

PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR  
DILEPTON MASS RESONANCES IN  $H \rightarrow 4\ell$  DECAYS USING THE CMS DETECTOR AT  
THE LHC

By

JAKE ROSENZWEIG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2022

© 2022 Jake Rosenzweig

This work is dedicated to the living and loving memory of Jacob Myhre.

## ACKNOWLEDGEMENTS

Without so many two- and four-legged blessings along the way, I could never have made it to this point in my academic career. Thus, I begin by giving my endless gratitude to my high energy physics mentors, Professors Andrey Korytov and Guenkah Mitselmakher, for granting me this one-of-a-kind opportunity to do real *science* at CERN.

To my wife, Suzanne Rosenzweig, for showing me that dreams *do* come true. To my mother and father, Vicki and John, who always reassured me that I could achieve anything I put my mind to. Sleep peacefully, Mom. To my siblings, Alex, Ryan, Devin, Jace, and Claudia who frequently and gently reminded me that life existed outside of grad school. To Auntie Rachel and Uncle Yuri, who just *get* me and have given to me everything I could have possibly asked for. To my mentor, Sheldon Friedman, and his wife, Rita Friedman (Rosenzweig), who chose to invest in my success at a young age. I have only made it this far thanks to their undying encouragement, love, and optimism. To Sheldon's best friend, Dr. Bernard Khoury, whose reputation and has helped pave my own path. To the many moms who generously gave unconditional support during the darkest of times and unequivocal love during the brightest times: Cyndi Reilly-Rogers, Dawn Hood, Margaret Sherrill, and Silet Wiley.

To Dr. Filippo Errico for his focus, leadership, selflessness, and patience in leading the Higgs mass analysis. To Dr. Lucien Lo for showing me the simplicity and beauty of Python in his typical laid-back way and for leading the dilepton analysis. To Dr. Noah Steinberg for showing me how majestically physics can be communicated from mind to mind. To Dr. Darin Acosta, for spending many hours of physics discussion with me and the other students, who have helped us build our "CMS Office Hours" Ari Gonzalez, Cris Caballeros, Jeremiah Anglin, Sean Kent, Evan Koenig, Neha Rawal, Nik Menendez, John Rötter. To the gents who paved my way to and through the world of CMS, Brendan Regnery and Bhargav Joshi. To my mentee, Matthew Dittrich, for accepting the baton of knowledge and making everything come full circle.

To my comrades for showing me what it takes to survive the core courses, Dr. Atool Divakarla, Dr. Brien O'Brendan, Dr. Donyell Guerrero, and Dr. Vladinar Martinez. To my quondam friend, Adamya Goyal, for his gentleness, humility, patience, and tutorage. To my Polish roommates in Saint-Genis-Pouilly for showing me what home away from home feels like:

Bartoszek, Dziadziuś, Karolina, and Sandruša. To the boys who have been there since the beginning: Jish, Willis, The Shane, Zacman, Duck, and Marcus for their clever competition and continual camaraderie which has shaped me to this day.

To Big Tree who stood as a symbol of strength, beauty, and life for centuries before us. As Irma's wild whirring winds worsened, the cacophony of ripping roots resounded throughout the western corridor. There I stood in that frozen moment—awestruck, speechless—watching her *fall* helplessly towards the physics building. What could have been a catastrophe of cataclysmic proportions was instead a gentle grazing against the north windows of NPB where, there, she was gracefully laid to rest.

Finally, to Existence, for this unpredictable, unbelievable blip of an experience called Life.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGEMENTS .....	4
LIST OF FIGURES.....	8
ABSTRACT.....	9
CHAPTER	
1    INTRODUCTION .....	10
2    THE LARGE HADRON COLLIDER.....	14
2.1  Motivation .....	14
2.2  The LHC at CERN.....	14
2.3  The Journey of a Proton at the LHC.....	16
3    THE CMS DETECTOR.....	20
3.1  The Silicon Tracker .....	21
3.1.1  The Pixel Detector.....	22
3.1.2  The Strip Detector .....	22
3.2  The Calorimeters .....	23
3.2.1  Electromagnetic Calorimeter.....	23
3.2.2  Hadron Calorimeter.....	24
3.3  The Solenoid and the Steel Return Yoke.....	26
4    HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^* \rightarrow 4\ell$ CHANNEL .....	33
4.1  Motivation .....	33
4.2  Analysis Overview.....	33
4.3  Background Estimation .....	38
4.3.1  Irreducible Background .....	38
4.3.2  Reducible Background .....	38
5    CONCLUSION.....	56
REFERENCES .....	56
BIOGRAPHICAL SKETCH .....	58

## LIST OF FIGURES

<u>Figures</u>	<u>page</u>
1-1 The elementary particles described by the SM. ....	10
1-2 Theoretical stability regions of the Universe. ....	12
2-1 (Left) A map image of the LHC ring. (Right) The CERN accelerator complex .....	15
2-2 Cross section of an LHC dipole magnet.....	18
2-3 Distribution of the average pile up during the LHC 2018 run.....	19
2-4 Diagram showing the constituents ( <i>partons</i> ) of two protons interacting in a pp collision... .	19
3-1 Life-size poster of the CMS detector.....	20
3-2 Points of interest along the LHC (Points 1–8). ....	28
3-3 Cut out of the CMS detector showing its various subdetector components. ....	28
3-4 A transverse view of CMS showing particle “filtration”.....	29
3-5 (Left) A simulation of the silicon tracker. (Right) The real silicon tracker within CMS....	30
3-6 A transverse view of the silicon pixel and strip detectors, explicitly labelling the different layers involved. ....	30
3-7 (Left) Cross-sectional view of the ECAL. (Right) An ECAL Dee. ....	31
3-8 (Left) ECAL crystals are grown in a lab. (Right) A single crystal. ....	31
3-9 Cross-sectional quadrant view of the HCAL components. ....	32
3-10 Cross section of CMS showing the magnetic field line strengths. ....	32
4-1 Branching ratios of Higgs boson decays as a function of the Higgs boson mass. ....	34
4-2 Feynman diagrams for the $gg \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$ and $gg \rightarrow ZZ^* \rightarrow 4\ell$ processes.....	36
4-3 Distribution of $m_{4\ell}$ from $H \rightarrow ZZ^* \rightarrow 4\ell$ events using Full Run 2 data.....	37
4-4 Diagram showing how RB prompt-lepton processes contribute to observed PTS-lepton events.....	42
4-5 Distributions of the predicted and observed RB yields for the 2P2F CR (2016 pre-VFP)... .	44
4-6 Distributions of the predicted and observed RB yields for the 2P2F CR (2016 post-VFP).. .	45
4-7 Distributions of the predicted and observed RB yields for the 2P2F CR (2017).....	46
4-8 Distributions of the predicted and observed RB yields for the 2P2F CR (2018).....	47
4-9 Distributions of the predicted and observed RB yields for the 3P1F CR (2016 pre-VFP)... .	48
4-10 Distributions of the predicted and observed RB yields for the 3P1F CR (2016 post-VFP).. .	49
4-11 Distributions of the predicted and observed RB yields for the 3P1F CR (2017).....	50

4-12 Distributions of the predicted and observed RB yields for the 3P1F CR (2018).....	51
4-13 Electron misidentification rates vs. electron $p_T$ for the OS Method. ....	54
4-14 Muon misidentification rates vs. muon $p_T$ for the OS Method. ....	55

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR  
DILEPTON MASS RESONANCES IN  $H \rightarrow 4\ell$  DECAYS USING THE CMS DETECTOR AT  
THE LHC

By

Jake Rosenzweig

August 2022

Chair: Andrey Korytov  
Co-Chair: Guenakh Mitselmakher  
Major: Physics

The mass of the Higgs boson is measured in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  ( $\ell = e, \mu$ ) decay channel and is found to be  $m_H = 125.38 \pm 0.11$  GeV; the most precise measurement of  $m_H$  in the world to date. The data for the measurement were produced from proton-proton (pp) collisions at the Large Hadron Collider with a center-of-mass energy of 13 TeV during Run 2 (2016–2018), corresponding to an integrated luminosity of  $137.1\text{fb}^{-1}$ , and were collected by the Compact Muon Solenoid experiment. This measurement uses an improved analysis technique in which the final state muon tracks are constrained to originate from the primary pp vertex. Using data sets from the same run, a search for low-mass dilepton resonances in Higgs boson decays to the  $4\ell$  final state is also conducted. No significant deviation from the Standard Model prediction is observed.

## CHAPTER 1 INTRODUCTION

The Universe, while overwhelmingly vast, is built from a remarkably few kinds of elementary (i.e., indivisible) particles. As shown in Fig. 1-1, any elementary particle can be classified into 1 of these 3 categories: matter particles (*fermions*), force-carrying particles (*gauge bosons*), and Higgs bosons. There are 12 kinds of fermions which can be split evenly into 2 groups, depending on with which forces they interact: those that interact via the electromagnetic (EM) force and the weak nuclear force are classified as *leptons*, of which there are 6 kinds (“*flavors*”), whereas those that interact via the EM, weak, *and* strong nuclear forces are classified as *quarks*, of which there are also 6 flavors. There are 4 kinds of gauge bosons, each of which is a force carrier for a specific force (the gluon is said to *mediate* the strong force, the photon mediates the EM force, while both the  $W^\pm$  and Z bosons mediate the weak force). Thus, all the diversity and manifestations of reality come from only 17 kinds of “building blocks”. It is the mission of particle physicists to understand the underlying mathematical structure that describes Nature in as accurate—and hopefully *concise*—a theory as possible. The best theory that stands today is called the Standard Model (SM) of particle physics.

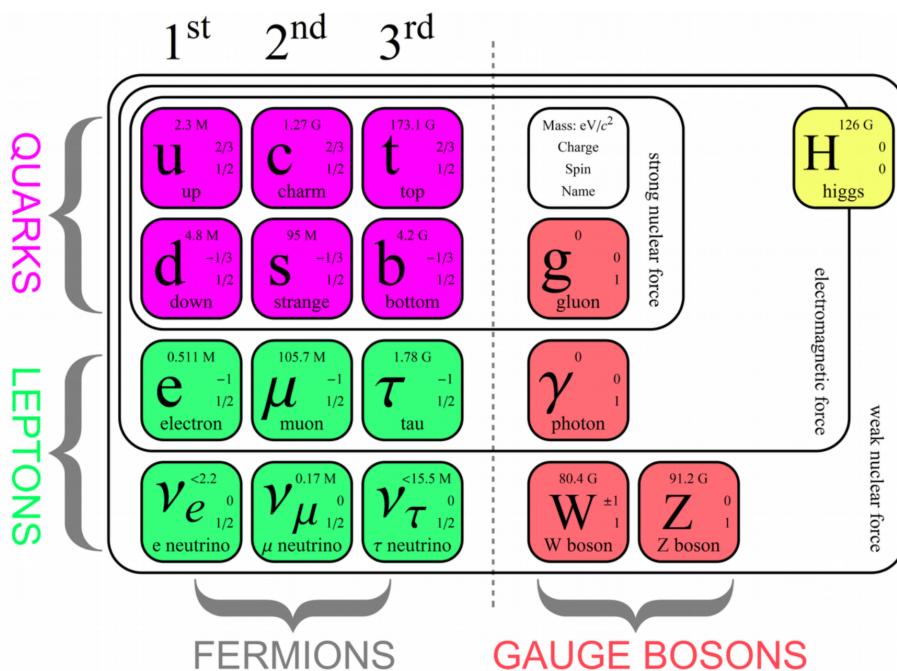


Figure 1-1. The elementary particles described by the SM.

The SM has bore witness to many triumphs over its approximately 70-year development: it predicted the existence of quarks which were experimentally confirmed in the mid-1970s; it predicted the existence of the tau neutrino which was experimentally confirmed in 2000; its most groundbreaking prediction was experimentally confirmed on July 4, 2012—almost exactly 10 years ago from the date of this dissertation writing—when the ATLAS and CMS collaborations announced the discovery of the Higgs boson [1–3].

Quantum Field Theory is the mathematical and conceptual backbone of the SM. Within its framework, all particles are *excitations* of their corresponding field so, e.g., *every* electron in the Universe is thought to be an excitation of the single electron field that permeates all of spacetime. The existence of the Higgs boson ( $H$ ) suggests that its corresponding field exists—the *Higgs field*—and, thus,  $H$  is the quantum excitation of that field. This all-pervasive Higgs field and its corresponding boson are generated mathematically via the Brout-Englert-Higgs (BEH) mechanism. Most SM particles—except for neutrinos, photons, and gluons—“acquire” their mass by interacting with the Higgs field. This is also how the Higgs boson itself acquires its mass ( $m_H$ ); by interacting with its own field!

The masses of SM particles *depend* on the value of  $m_H$  . . . so what is its value? Unfortunately,  $m_H$  is a free parameter of the SM, so theory is unable to provide a value for  $m_H$  based solely on other fundamental constants. Instead, the value of  $m_H$  must be measured by experiment—and has been measured multiple times [4–6]. Although  $m_H$  has already been measured, it is important to continually improve the measurement by lowering the uncertainties on (i.e., increasing the precision on) the mass value. A more precise value of  $m_H$  is two-fold: first, it improves the theoretical limits on the masses of the other elementary particles and, second, it very well may determine the stability of our universe, as shown in Fig. 1-2.

The production of a Higgs boson is only feasible at conditions close to those thought to exist at the beginning of the Universe. This grand achievement is accomplished frequently by the Large Hadron Collider (LHC), located on the border of France and Switzerland. It is the largest and most powerful proton-proton (pp) collider ever made. The LHC accelerates protons to incredible

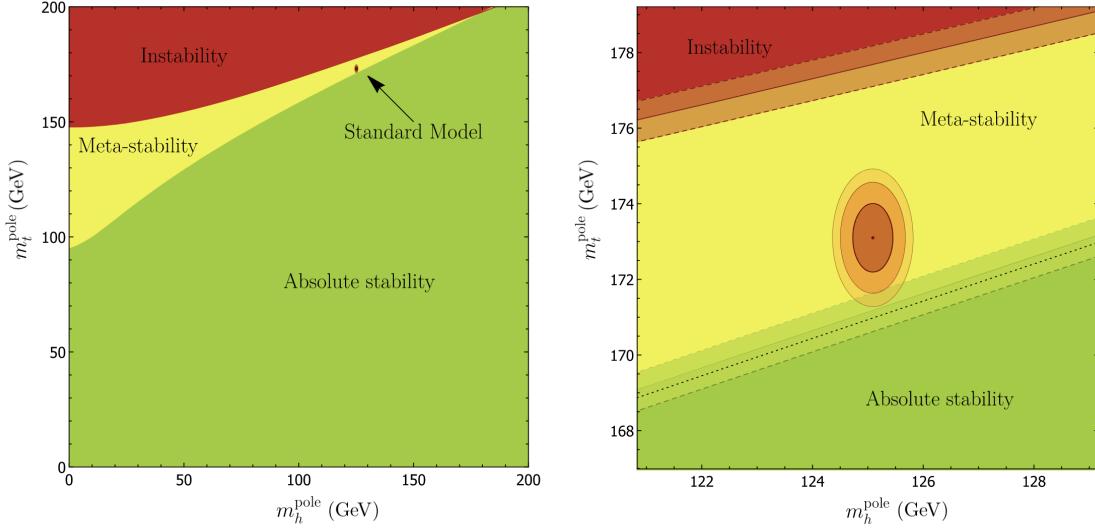


Figure 1-2. (Left) Theoretical stability regions of the Universe based on the pole masses of the top quark ( $m_t^{\text{pole}}$ ) and Higgs boson ( $m_h^{\text{pole}}$ ). (Right) A closeup of the SM region of the left plot. The contours represent the 68%, 95%, and 99% confidence levels based on the experimental uncertainties of  $m_t^{\text{pole}}$  and  $m_h^{\text{pole}}$ . Plots taken from [7] and units added to all axes.

speeds, very close to the speed of light. When these fast-moving protons collide, the pp collisions can have a center-of-mass energy as high as 13 TeV. Newly produced particles spew out of the collision points and are analyzed by detectors like the aforementioned ATLAS and CMS experiments. These enormous particle detectors have thousands of scientists performing dozens of analyses to look for hints of beyond Standard Model (BSM) physics, extra dimensions, miniature black holes, and more.

This dissertation utilizes data collected by the CMS experiment during the LHC Run 2 (2016–2018) to perform the world’s best precision measurement of  $m_H$  to date. This new measurement utilizes the following improvements compared to previous measurements:

- Nearly four times as much collected data from Run 2 ( $\mathcal{L}_{\text{int}} = 137.1 \text{ fb}^{-1}$ ) vs. the data used for the 2016 measurement ( $\mathcal{L}_{\text{int}} = 35.9 \text{ fb}^{-1}$ ).
- Four final-state categories:  $4\mu$ ,  $4e$ ,  $2e2\mu$ ,  $2\mu2e$ . In previous measurements, the last two final states (the mixed-flavor states) were combined, when truly they have different kinematical properties (depending on into which lepton pair the  $Z_1$  decayed): different peak widths (instrumental resolutions), different signal efficiencies, and different relative levels of

reducible background.

- Ultra-Legacy (UL) reconstruction for muon, electron, photon, and jet tracks. This significantly improves electron momenta and improves the other particle momenta, though to a lesser degree.
- The measurements of muon  $p_T$  are improved by constraining the muon tracks to originate from the interaction vertex (also called a *vertex constraint*).
- When extracting the value of  $m_H$  in past measurements, a 3D pdf  $(m_{4\ell}, \sigma_{m_{4\ell}}, \mathcal{D}_{\text{bkg}}^{\text{kin}})$  was built into a factorized form  $f(m_{4\ell}, \sigma_{m_{4\ell}} | m_H) \cdot g(\mathcal{D}_{\text{bkg}}^{\text{kin}} | m_{4\ell})$ , which was later found to contain an existing correlation between  $\sigma_{m_{4\ell}}$  and  $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ . To account for this correlation, now the events are split into 9 categories based on the per-event *relative* mass uncertainty  $\left(\frac{\sigma_{m_{4\ell}}}{m_{4\ell}}\right)$  and, for each, a 2D pdf  $(m_{4\ell}, \mathcal{D}_{\text{bkg}}^{\text{kin}} | m_H)$  is built.
- The systematic uncertainties on electron and muon momentum scales ( $p_T^{e,\mu}$ ) are reduced, thanks to a more detailed analysis on the uncertainties. This has the additional effect of significantly reducing the uncertainty on the per-event four-lepton mass resolution.

The following chapters of this dissertation begin by describing the function and engineering of the Large Hadron Collider in Chapter 2. Then, a thorough description of the CMS Experiment and its composite subdetectors is given in Chapter 3. Next, the details of the precision measurement of the Higgs boson mass using the LHC Run 2 data is discussed in Chapter 4. Finally, the results of the Higgs boson mass measurement analysis is summarized in Chapter 5.

## CHAPTER 2

### THE LARGE HADRON COLLIDER

#### 2.1 Motivation

Although the SM has shown to be an astoundingly accurate framework so far, it must continue to be scrutinized by the barrage of measurements that either confirm or contradict its predictions. Interestingly, a recent measurement of the mass of the W boson has shown significant deviation from SM predictions, with a sensitivity of  $7\sigma$  [8]. After all, undeniable fact comes from the reproducible results obtained from *measurement*—not from some theoretical model which *may* or *may not* describe reality. Whenever the predictions of a model directly contradict the results from measurement, the model must necessarily be cast aside and replaced by one whose predictions concur with the results of measurement.

So how *are* measurements obtained in the realm of particle physics? Modern day physicists study the fundamental constituents of matter and their interactions by using state-of-the-art technologies combined with time-tested methodologies: by smashing tiny bits of matter together to turn them into even *tinier* bits. Such is the purpose of the world’s largest and most powerful particle accelerator—the Large Hadron Collider (LHC).

#### 2.2 The LHC at CERN

Deep beneath the surface of the earth (50–175 m), the LHC straddles the border shared by France and Switzerland. Sandwiched between the scenic Jura mountains to the northwest and the sprawling city of Geneva (French: Genève) to the southeast is CERN. To illustrate the enormous circumference (26.659 Km) the sheer size of this circular accelerator, Fig. 2-1 (Left) shows the LHC drawn on a map. For reference, the inscribed area of the LHC ( $56.7 \text{ Km}^2$ ) is almost four times greater than the area of the neighboring city of Geneva ( $15.9 \text{ Km}^2$ ).

The LHC is not only a particle *accelerator* but also a proton-proton (pp), proton-lead ion, and lead-lead ion *collider*. By sending one particle beam clockwise around the ring and the other beam counterclockwise, the charged particles are carefully maneuvered around the ring using dipole magnets and collimated into tight proton *bunches* using quadrupole magnets before they ultimately collide at 4 specific points along the LHC, as shown in Fig. 2-1 (Left, red stars). When the LHC is fully powered, *each* proton in the beam carries an average energy of 6.5 TeV which

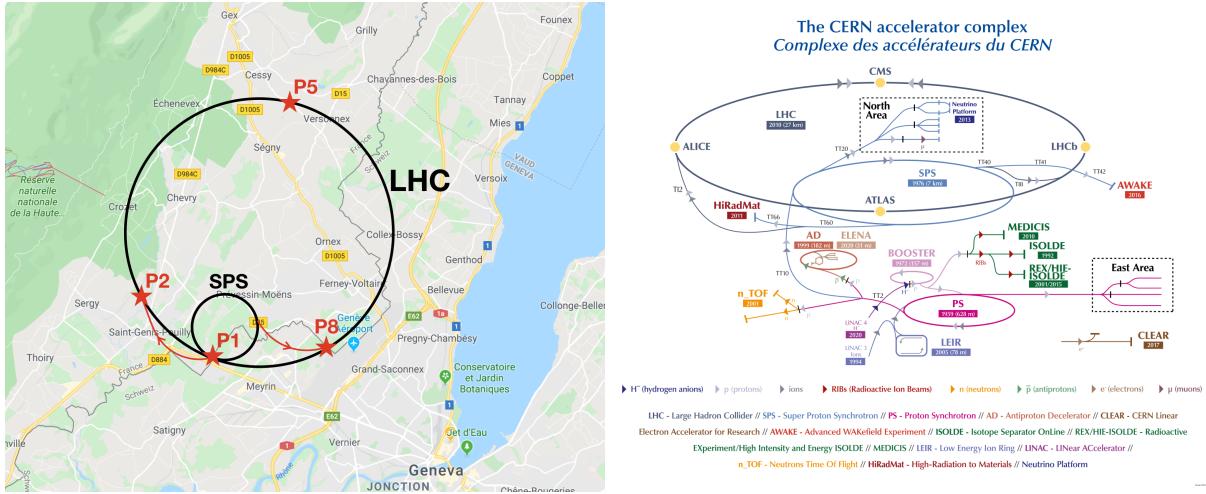


Figure 2-1. (Left) The LHC ring (bigger ring) and the Super Proton Synchrotron (smaller ring) with the nearby town of Geneva for size comparison. The four red stars indicate the pp collision points. (Right) The accelerator complex at CERN.

gives a single pp collision a center-of-mass energy of 13 TeV. This emulates the conditions theorized to exist at the beginning of the universe, which allows cosmological studies to be carried out. The hugely energetic pp collisions cause the quark and gluon constituents within the protons to interact with each other and transform into new particles. The newly created particles and the residual particle debris are ejected away from the collision point—whether straight down the beampipe, completely orthogonal to it, or somewhere in between. A massive particle detector is stationed at each of the 4 collision points to detect the outgoing particle “spray”. The 4 main particle detectors and their locations along the LHC are:

- A Toroidal LHC ApparatuS (ATLAS)—located at the first collision point (P1),
- A Large Ion Collider Experiment (ALICE)—located at P2,
- the Compact Muon Solenoid (CMS, Chapter 3) experiment—located at P5, and
- the LHC-beauty (LHCb) experiment—located at P8.

The world-renowned feat of digging the tunnel for, constructing, commissioning, and monitoring the LHC was made possible by CERN: the European Organization for Nuclear Research (French: *Conseil Européen pour la Recherche Nucléaire*). CERN is an international collaboration of—at the time of this writing—more than 33 countries, each of which is considered

either a Member State, an Associate Member State, or an Observer. The complex of CERN (Fig. 2-1, Right) is located just to the west of P1 and is akin to a small science *city*—complete with many offices, manufacturing facilities, and experiments such as the Antiproton Decelerator (AD), the Neutrons Time of Flight (n\_TOF), and the Isotope Separator OnLine (ISOLDE) experiments. Although the LHC is the most famous of the accelerators at CERN, its fame is only made possible by a series of smaller and lesser-known accelerators that *feed* the LHC. Therefore, a natural way to explore the intricacies and inner workings of the LHC is to follow the path of one of its “inhabitants”—a single proton—as it makes its way to and through the gigantic collider.

### 2.3 The Journey of a Proton at the LHC

The journey to discovery begins in a surprisingly small tank of hydrogen gas ( $H_2$ ) located in the LINAC4 building at the main CERN site. Inside this tank, a proton—conveniently called P. Roton—and approximately  $10^{23}$  other protons coexist in bound states as molecules of  $H_2$ . Although the tank has a meager mass of 10 Kg, it has enough protons inside to keep the LHC colliding protons for over *200 000 years* of constant operation.

Protons get injected into a series of increasingly larger accelerators. It begins with the *Linac4*—a linear particle accelerator—that accelerates hydride ions ( $H^-$ ) to 160 MeV which eventually make their way into the *Proton Synchrotron Booster* (PSB). In the PSB, each hydride ion has its electron pair completely stripped away, leaving only the bare proton. The protons then enter a series of circular accelerators, each machine feeding protons into the next while increasing the proton center-of-mass energy by at least 1 order of magnitude.

- Protons are then accelerated to 2 GeV at which point they are injected into the *Proton Synchrotron* (PS).
- The PS then increases the proton energy to 26 GeV to be fed into the *Super Proton Synchrotron* (SPS).
- The penultimate step is for the SPS to further energize the protons to a center-of-mass energy of 450 GeV.
- Finally the protons enter the LHC and are further accelerated to the maximum energy of 6.5 TeV using RF cavities to kick them.

The protons would travel in a straight line if not for the 1232 dipole magnets sprinkled along the LHC. They are made out of copper-clad niobium-titanium, require the cryogenics of 96t of superfluid helium-4 to ultimately turn the proton beams. There are also 506 quadrupole magnets that compress the proton bunches to make them as linear as possible before they collide. Finally two incoming proton bunches approach a common collision point. Beam pipes are conjoined in an “X” shape where 2 proton bunches cross—a *bunch crossing* (BX). Out of more than 40 million pp collisions that could have occurred, a mere 50 collisions take place on average (i.e., only 0.0001%). This is a testament to just how small protons truly are.

Just as the PS feeds protons into the SPS, which feeds protons into the LHC, so too is it being considered for the LHC to feed a new project—the 100 Km Future Circular Collider.

In the event that P. Roton does not collide with any of the oncoming protons, then it is simply “recycled” and continues going around the LHC ring for another opportunity at a pp collision.

Contrary to what some people may think, protons are not sent one by one at each other, hoping for a collision. Instead 100 billion protons are packed together into a “proton bunch”. A single proton bunch is about the size of a human hair ( $\approx 50\text{ }\mu\text{m}$  wide and  $\approx 10\text{ cm}$  long). The clockwise and counterclockwise rings are filled to a maximum of 2808 proton bunches, each one spaced 25 ns apart, and then sent to collide.

It requires an incredibly strong magnetic field to turn the protons as they make their revolutions around the LHC. Recall that charged particles bend in a magnetic field, via the Lorentz force. Therefore, the LHC is equipped with 1232 dipole magnets distributed all along the length of the beam pipe to keep the proton bunches turning in the tunnel. The cross section of such a dipole magnet is shown in Fig. 2-2. Each dipole magnet is 14.3 m long, weighs 35 t, cost nearly 500 KCHF to produce, and has nearly 11 700 amps of current running through it. Only with such massive currents is it possible to generate the appropriate magnetic field strength of 8 T to keep the protons turning. The magnetic field is maintained by titanium-niobium coils, which are kept under cryogenic conditions using liquid helium to achieve the necessary temperature of 1.9 K to reach a superconducting state; this temperature is colder than that of outer space.

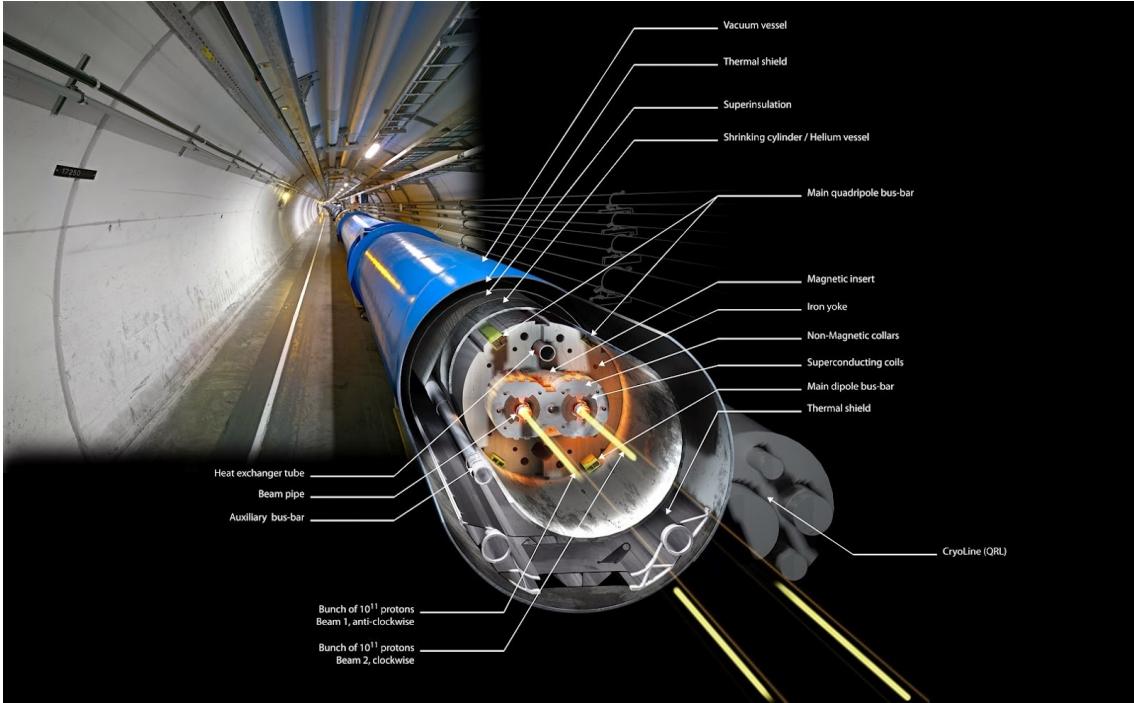


Figure 2-2. A cross section of one of the 1232 dipole magnets which span the entire length of the LHC tunnel.

There are only four specific “Points” along the LHC where the proton bunches actually cross, as shown in Fig. 2-1 (Left, red stars). At each of these four points, there is a unique and gigantic particle detector to catch all the decay products from the pp collisions.

As the two bunches are just about to cross one another, they are squeezed down using quadrupole magnets, focusing the beams more tightly, increasing their chance for tasty pp collisions. During such a bunch crossing (BX), amazingly most of the protons just pass right by one another; out of the possible 100 billion possible collisions that could have occurred, Fig. 2-3 shows that on average only 32 collisions occurred per BX in the LHC 2018 run, according to a particle detector called CMS, described in Chapter 3. It should be mentioned that the luminosity of the LHC is on the order of  $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ .

As the proton bunches whiz around the LHC, they are given “kicks” from radio-frequency (RF) cavities, which accelerate the protons to a max speed of 99.999 996% $c$ . It is analogous to the timing required when pushing someone on a swing: push at just the right time to increase their momentum. At this speed, *each proton* carries 6.5 TeV of energy, such that a single pp collision

contains a monstrous center-of-mass energy of  $\sqrt{s} = 13 \text{ TeV}$ —more than enough energy to create new particles like top quarks, Higgs bosons, and potentially BSM particles (Fig. 2-4). In order to analyze such interesting particles, one needs to detect the outgoing particles produced from pp collisions. One such dedicated particle detector is located at Point 5—the Compact Muon Solenoid (CMS) detector—and will be described in the following chapter.

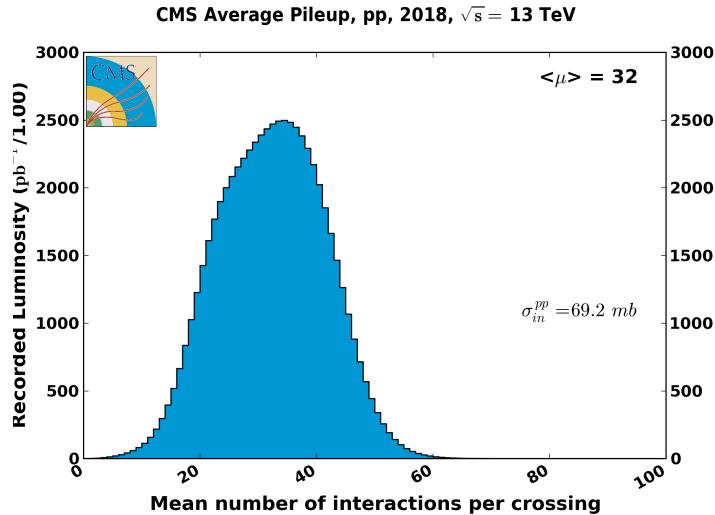


Figure 2-3. Histogram showing the distribution of the average number of pp collisions per proton bunch crossing (pile up) which CMS recorded during the LHC 2018 run.

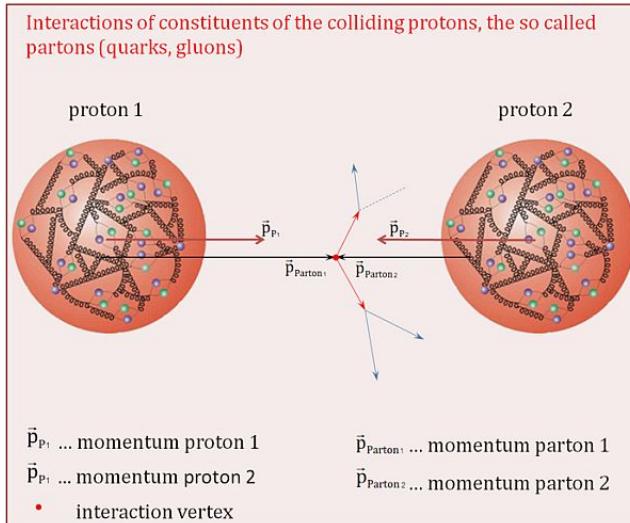


Figure 2-4. Two protons can be smashed together at very high energies to have their constituent partons interact and convert their high energies into new kinds of matter.

## CHAPTER 3

### THE CMS DETECTOR



Figure 3-1. Life-size poster of the CMS detector, taken during CERN Open Days 2019 in the SX5 warehouse where parts of CMS were assembled.

Weighing in at 14 000 tonnes, standing 15 m (5 stories) tall, and reaching 29 m long, the Compact Muon Solenoid (CMS) experiment is one of the four major particle detectors at the LHC (Fig. 3-1). As shown in Fig. 3-2, CMS is situated approximately 100 m under the earth at the fifth collision point (*Point 5*) along the LHC. In 2012, both CMS and its competing experiment, A Toroidal LHC ApparatuS (ATLAS), independently discovered the Higgs boson.

Recall that the LHC (Chapter 2) directs two proton bunches together on a collision course every 25 ns to produce thousands of new particles per pp collision. These outward-moving particles travel all sorts of directions away from the interaction point. CMS is built around the interaction point in a series of cylindrical subdetectors for nearly hermetic coverage so that most of the particles *must* travel through CMS. The detector sports a solenoid—for which CMS was named—which generates a 3.8 T uniform magnetic field that points longitudinally down the central axis of CMS. This strong magnetic field applies a Lorentz force on the outgoing charged particles, causing them to follow helical, momentum-dependent trajectories. These curved tracks separate from one another which assists in particle identification. Electrically neutral particles experience no Lorentz force and thus travel in straight lines.

The subdetectors measure the properties of the outgoing particles and carefully filter them

out in a clever way (Fig. 3-3). As particles pass through the subdetectors, they interact in a variety of ways. For example, charged particles can leave “hits” in the silicon tracker system and continue through to the next radially outward subdetector, whereas hadronic matter tends to be captured by the hadronic calorimeter. Generally, hits can be reconstructed into particle tracks. From the track curvature, the charge and momentum of the particles can be deduced. Depending on *which* subdetector (or combination of subdetectors) produced signals, the type of particle can be deduced. A few example particles and their associated tracks are shown in Fig. 3-4.

Before discussing each subdetector in the following sections, it is useful to define the coordinate system used in CMS; a typical, right-handed, three-dimensional Cartesian coordinate system ( $x, y, z$ ) is chosen, whose center  $(0, 0, 0)$  is placed at the nominal pp collision point within CMS. The  $x$  axis points towards the center of the LHC, the  $y$  axis points vertically upward, and the  $z$  axis points westward towards the Jura mountains, tangential to the beam direction. Since CMS covers almost the entire spherical  $4\pi$  steradians around the interaction point, it is convenient to use spherical coordinates  $(r, \phi, \theta)$ , in which  $r$  measures the radial distance in the  $x$ - $y$  plane,  $\phi$  measures the azimuthal angle in the  $x$ - $y$  plane as measured from the  $x$  axis, and  $\theta$  measures the polar angle as measured from the  $z$  axis. When dealing with ultra-relativistic particles like those produced at the LHC, special-relativistic effects like length contraction must be taken into account and so the coordinate  $\theta$  becomes frame-dependent. It is thus helpful to convert  $\theta$  to the Lorentz-invariant quantity called pseudorapidity ( $\eta$ ), defined as  $\eta = -\ln[\tan(\theta/2)]$ .

In the remainder of this chapter, the subdetectors of CMS are described in detail: the silicon tracker in Sec. 3.1, followed by the electromagnetic and hadron calorimeters in Sec. 3.2, then the solenoid and yoke system in Sec. 3.3.

### 3.1 The Silicon Tracker

At the heart of CMS is one of the world’s largest silicon detectors: the silicon tracker. The main goal of the silicon tracker is not to capture outgoing particles but to very precisely measure the hits from the charged particles as they pass through it. The tracker also assists in vertex identification, differentiating between primary and secondary vertices, the latter of which often

comes from B meson decays. When multiple pp collisions occur within the same BX (so-called *pile up*), the tracker distinguishes between proton collisions with a resolution of about  $100\text{ }\mu\text{m}$  longitudinally and  $50\text{ }\mu\text{m}$  transverse to the beam pipe. This is crucial to resolve which outgoing particles came from which pp vertex.

The tracker consists of two types of pure silicon detectors: the pixel detector and the strip detector, each of which is described in turn below.

### 3.1.1 The Pixel Detector

The innermost part of the silicon tracker is the pixel detector, which is the closest subdetector to the interaction point. The pixel detector is composed of 66 million silicon “pixels”, as shown in Fig. 3-5 (Left, pink). A single pixel is  $100\text{ }\mu\text{m} \times 150\text{ }\mu\text{m}$  and, collectively, they cover a sensitive area of  $1.9\text{ m}^2$ . Because it sits only 8 cm away from the beam pipe, the pixel detector receives the highest particle flux than any other subdetector: around 10 million particles/ $\text{cm}^2$  per second.

The pixel detector is made of three cylindrical layers and two endcaps that surround the beam pipe. In total, the pixel detector has around 6,000 connections (channels) per  $\text{cm}^2$ .

After the LHC Run 1 was completed, the accelerator received luminosity upgrades during the 2013–2014 Long Shutdown. To handle these higher luminosities, the pixel detector was replaced by the CMS Phase-1 pixel detector during the LHC technical stop in 2016–2017. The upgrades outfitted the detector with four barrel layers and three endcap disks per side, which allowed for particle detection up to  $|\eta| < 2.5$ . The overall mass of the pixel detector decreased and granted the detector with better tracking capability.

### 3.1.2 The Strip Detector

The outer part of the silicon tracker is called the strip detector, which has 10 million detector strips spread across 10 cylindrical layers. The first 4 layers belong to the tracker inner barrel (TIB) and the remaining 6 layers belong to the tracker outer barrel (TOB), Fig. 3-5 (Left, green and blue, respectively). Both the TIB and TOB have two endcaps associated with them, the TID and TEC, respectively. Accounting for all of its components, the strip detector is sensitive to  $200\text{ m}^2$ .

Fig. 3-6 gives a clearly labelled transverse illustration of the pixel and strip detectors.

## 3.2 The Calorimeters

### 3.2.1 Electromagnetic Calorimeter

**Overview:** Particles that pass through the silicon tracker 3.1 encounter the electromagnetic calorimeter (ECAL). Those particles which interact electromagnetically but not strongly (i.e., mostly photons and electrons) are typically absorbed by the ECAL. The particle's energy is then transferred to the ECAL in the form of an electromagnetic (EM) shower. The size and shape of the EM shower provide information about the particle's energy and trajectory. Since the Higgs boson can decay into two photons, the ECAL played a critical role in detecting this decay mode.

**Design:** The ECAL is a hermetic, cylindrical, homogeneous sub-detector that consists of a barrel (EB), two endcaps (EE), and a preshower detector in front of each endcap (Figure 3-7, Left). The EB covers  $|\eta| < 1.479$  while the EE covers  $1.479 < |\eta| < 3.0$ . The entire subdetector is composed of transparent lead tungstate ( $\text{PbWO}_4$ ) crystals that point axially towards the center (i.e., towards the interaction point) of CMS. The transparent crystals, one of which is shown in Fig. 3-8 (Left), have a high density ( $8.28 \text{ g/cm}^3$ ) which provides the ECAL with radiation resistance and a short radiation length ( $X_0 = 0.89 \text{ cm}$ ). Because so many crystals are used (61 200 crystals in the EB and 7324 in the EE), the ECAL has excellent energy resolution and fine granularity. Each endcap is composed of two *Dees*, one of which is shown in Figure 3-7 (Right). A single Dee carries 3662 crystals. Crystals in the barrel are tapered, having front-face dimensions  $2.2 \times 2.2 \text{ cm}^2$ , back-face dimensions  $2.6 \times 2.6 \text{ cm}^2$ , and are  $23.0 \text{ cm}$  long ( $25.8 X_0$ ). Crystals in the endcaps are also tapered, with front-face dimensions  $2.862 \times 2.862 \text{ cm}^2$ , back-face dimensions  $3.0 \times 3.0 \text{ cm}^2$ , and are  $22.0 \text{ cm}$  long ( $24.7 X_0$ ). This gives a single crystal from the barrel a volume of approximately  $132.5 \text{ cm}^3$  (i.e.,  $132.5 \text{ mL}$ )—about the volume of a small cup of coffee—yet it has a surprisingly hefty mass of  $1.5 \text{ Kg}$  [9].

**Physics:** When electrons or photons pass through the ECAL, they create an EM shower. Electrons radiate more photons as they accelerate around  $\text{PbWO}_4$  nuclei, in a process called *bremsstrahlung*. Meanwhile, near the presence of a nucleus, high-energy photons pair produce into  $e^+e^-$ . This cycle of electron/photon production disperses all the initial particle energy into a

spray of decreasingly lower-energy particles until it all runs out; this is the EM shower.

The ECAL crystals then scintillate (emits photons) in proportion to the amount of energy deposited by the interacting particle. The scintillator photons are detected by avalanche photodiodes on the back of each barrel crystal or by vacuum phototriodes in the endcap crystals (Fig. 3-8, Right). After 1 BX (25 ns), approximately 80% of the scintillated light is emitted.

An energy deposit in the ECAL could come from either an electron or a photon. In order to tell the difference, information from the silicon tracker is used. Charged particles, like electrons, will leave hits in the tracker and follow a curved path, whereas photons are electrically neutral and thus will not show any signs within the silicon tracker. So long as the tracker and ECAL communicate effectively with each other, then they help distinguish between electrons and photons. Charged hadrons interact only minimally with the ECAL, instead continuing on to the Hadron Calorimeter. Neutral hadrons can be detected by the ECAL preshower near the ECAL endcaps which helps distinguish a single photon from  $\pi^0$  mesons as they decay into two photons with a narrow opening angle, making it look as if the two photons are a single photon. The preshower detector allows CMS to distinguish between collimated, low-energy diphoton pairs and single high-energy photons.

From the original spray of particles that leave the interaction vertex, the short-lived particles have decayed into lighter, more stable, particles and the ECAL has filtered out most of the electrons and photons. The remainder of the spray is comprised of both hadronic matter and muons, however the hadrons are many times more numerous than the muons and so are filtered out next.

### 3.2.2 Hadron Calorimeter

**Overview:** The particles that survive the ECAL—typically only muons and hadrons—then enter the hadron calorimeter (HCAL). Its primary purpose is to absorb the hadronic matter emerging from the interaction point and to measure the corresponding jet energies. The absorbed jets cause the HCAL to scintillate photons which are then converted into electrical signals. These signals help deduce the original jet energies and missing transverse energy ( $E_T^{\text{miss}}$ ) from the event.

**Design:** Dissimilar to the ECAL (Sec. 3.2.1) in material composition but similar to it in shape, the HCAL is a brass cylindrical scintillator. Although it has a barrel (HB) and two endcaps (HE), it has two more detectors than the ECAL: the outer calorimeter (HO) and the forward calorimeter (HF). The HB spans the pseudorapidity range  $|\eta| < 1.3$ , the HE spans  $1.3 < |\eta| < 3$ , and the HF spans  $3 < |\eta| < 5.2$ , as shown in Fig. 3-9. With a thickness of over 1 m, the HB is sandwiched between the barrels of the ECAL and the solenoid (Sec. 3.3) at radial values  $r = 1.77$  m and  $r = 2.95$  m, respectively. Because the HB and HE are located within the solenoid's strong magnetic field of 3.8 T, they were both constructed out of a non-magnetic absorber called *C26000 cartridge brass*. This absorber has a density of  $8.53 \text{ g/cm}^3$  and an interaction length ( $\lambda_I$ ) of 16.42 cm. The thickness of the HB increases as  $1/\sin\theta$  so that at  $|\eta| = 0(1.3)$  the absorber thickness is  $5.82(10.6) \lambda_I$ . The HB is composed of two half-barrels, where each half-barrel is built from 18 identical azimuthal wedges and each wedge spans  $20^\circ$ . Each wedge is divided into four  $\phi$  segments so that a single  $\phi$  segment spans  $\Delta\phi = 0.087$ .

Since the volume available to the HCAL is so limited—and in order to stop any particles that might traverse the entire HCAL and solenoid—the HO (the *tail catcher*) is situated outside the barrel of the solenoid. The HF is located 11.2 m from the interaction point. All tiles within a single  $\phi$  segment are grouped together into a single tray unit. The scintillator is also segmented into 16  $\eta$  sectors, the first(last) of which is located at  $|\eta| = 0(1.3)$ . This way each tile covers  $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$ . Each layer has 108 trays.

**Physics:** Since hadrons are the only particles to interact via the strong force, the HCAL is designed to have a high nuclear density. This ensures ample opportunity for hadrons to radiate gluons and convert with the Similal to the ECAL, the HCAL will scintillate in proportion to the amount of energy of the captured particle. The incoming hadrons will *hadronize* (i.e., produce a hadronic shower), generating jets of quarks and gluons which are bound in various ways forming protons, neutrons, pions, kaons, etc. Interestingly, the HCAL is made using over a million old, brass shell casings from the Russian Navy back from World War II.

About 34% of the particles produced from LHC pp collisions enter the HE region, so the

HE was built to handle high rates (MHz).

The entire HCAL utilizes approximately 70,000 plastic scintillator tiles. The active material in the HB is 3.7-mm-thick Kuraray SCSN81 plastic scintillator, selected for its radiation hardness and long-term stability. Hadron showers  $\rightarrow$  tiles scintillate  $\rightarrow$  scintillated photons are collected by 0.94-mm-diameter green double-cladded wavelength shifting (WLS) fibers (Kuraray Y-11), which carry the light to hybrid photodiodes (HPD).

### 3.3 The Solenoid and the Steel Return Yoke

The Compact Muon *Solenoid* sports one of the world's most energetic solenoids which is paramount to the success of CMS. Particles that exit the HCAL (subsec. 3.2.2) arrive at the cylindrical magnet which is 12.5 m in length, has a bore diameter of 6 m (6.3 m when cold), and generates a uniform 3.8 T magnetic field parallel to the beam line. To produce such a large and uniform magnetic field inside the approximately  $360\text{m}^3$  volume (Fig. 3-10), an 18 000 amp current travels through the 4-layer, superconducting, NbTi coils. This magnetic field stores a massive 2.6 GJ of energy—approximately the kinetic energy of an Airbus A320 in flight—and is 100 000 times stronger than Earth's magnetic field, as measured on the surface. The magnet has such a large stored-energy-to-cold-mass ratio (11.6 KJ/Kg) that it experiences a physical deformation of 0.15% while being energized.

As charged particles travel through any magnetic field, they experience a magnetic (Lorentz) force perpendicular to their direction of travel. The Lorentz force ( $\vec{F}_B$ ) exerted on a particle with charge  $q$  depends on the particle's velocity ( $\vec{v}$ ) and the magnetic field ( $\vec{B}$ ) via the cross product

$$\vec{F}_B = q\vec{v} \times \vec{B}.$$

Since the force is necessarily perpendicular to the velocity, the resulting trajectory is helical. Projecting the helix on the  $x$ - $y$  plane (since the magnetic field points in the  $+z$  direction) allows the particle tracks to typically be separated from one another. Each track has a corresponding radius of curvature ( $R$ ) which relates to its transverse momentum ( $p_T$ ) through

$$p_T = qBR.$$

The relative change in  $p_T$  (i.e., the *momentum resolution*) is given by

$$\frac{\delta p_T}{p_T} \propto \frac{p_T}{BL^2}. \quad (3-1)$$

**Steel Return Yoke:** Most of the mass of CMS comes from the *steel return yoke* which helps to redirect the magnetic field back on itself. The yoke system constitutes 10 000 tonnes, which is 89% of the total mass of CMS. It is comprised of 5 wheels and 2 endcaps.

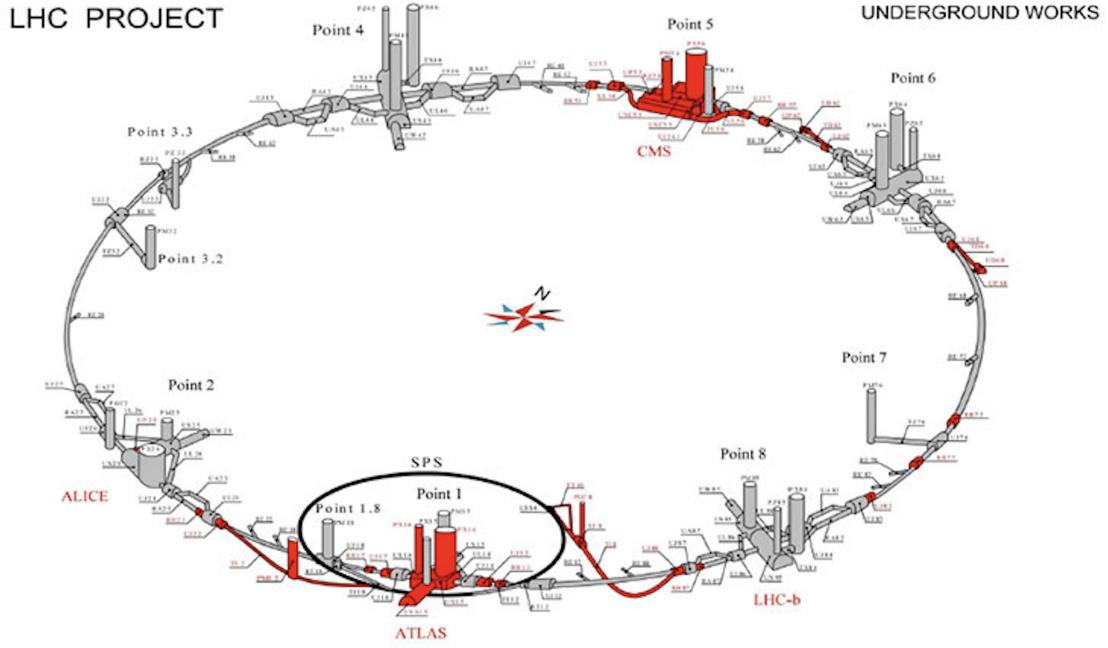


Figure 3-2. Points of interest along the LHC (Points 1–8). Collisions occur at Points 1 (ATLAS), 2 (ALICE), 5 (CMS), and 8 (LHCb), whereas the remaining points are used for LHC beam maintenance and testing.

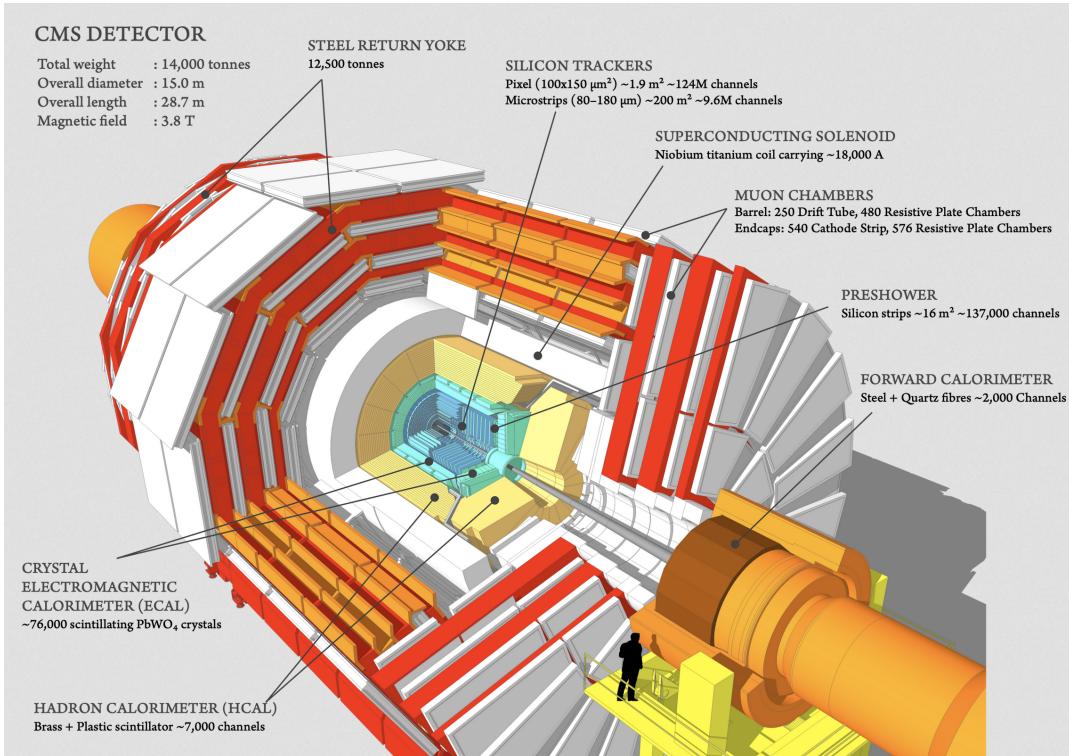


Figure 3-3. Cut out of the CMS detector showing its various subdetector components.

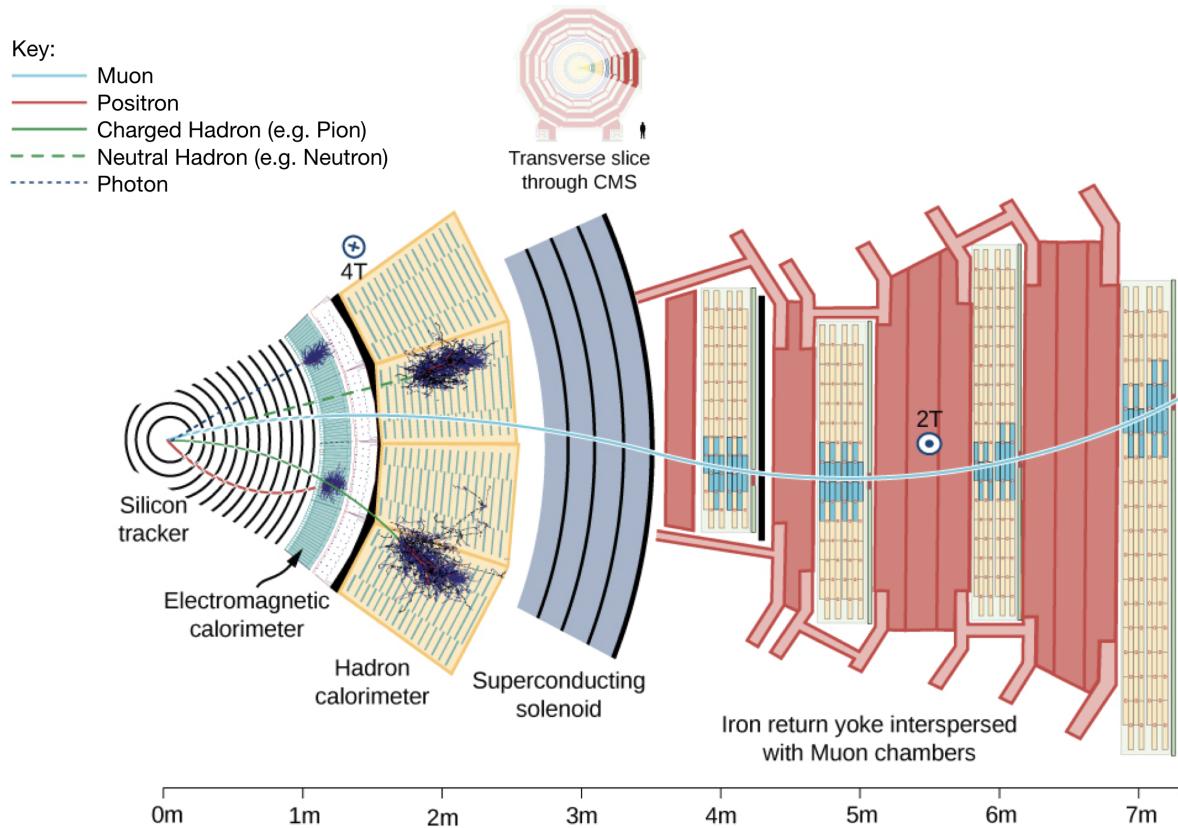


Figure 3-4. A transverse view of CMS showing the “filtration process” as different particles pass through different subdetectors. A positron (solid red line) curves due to the presence of the magnetic field and gets stopped in the ECAL, creating an EM shower. A photon (blue dashed line) does not get detected at all by the Silicon Tracker, since it has no electric charge. It continues through to the ECAL and makes a shower here, like the positron. Charged hadrons (solid green line) will show curved tracks from the Silicon Tracker, may leave some trace in the ECAL, but primarily get stopped by the HCAL creating hadronic showers. Neutral hadrons (dashed green line) do not interact with the tracker, and only undergo EM showers a little in the ECAL, but show most energy deposits in the HCAL. Muons (solid blue line) are detected by the Silicon Tracker and then mostly pass through the other subdetectors without interacting until they finally reach the Muon System. Using the Lorentz force law and knowing which direction the magnetic field is pointing, one can deduce the sign of the charge of the particle. Based on the radius of curvature from the trajectory, one can then calculate the momentum and energy of the particle.

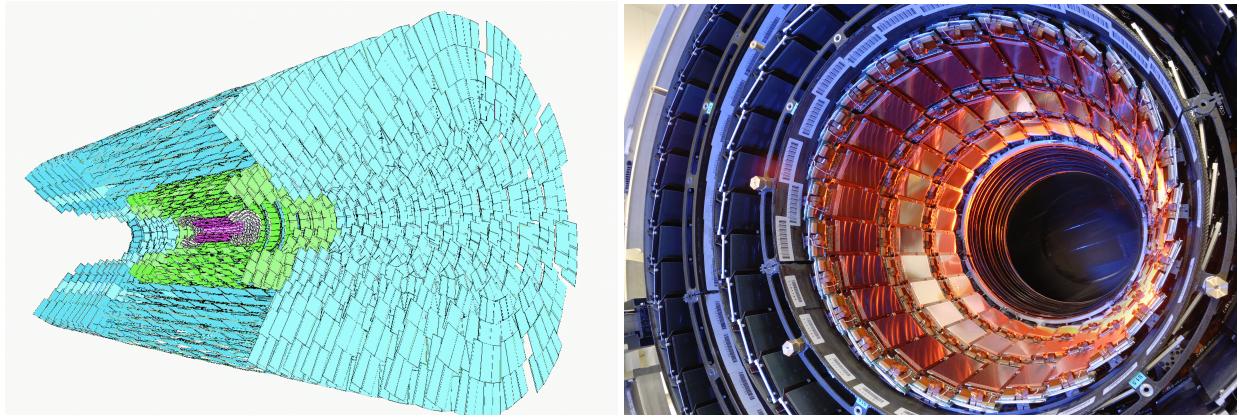


Figure 3-5. (Left) A simulation of the silicon tracker, showing the 3 cylindrical layers of the pixel detector (pink), 4 layers of the TIB (green), and the 6 layers of the TOB (blue) of the strip detector. The endcap components are also shown. (Right) A picture of the real silicon tracker at the center of CMS.

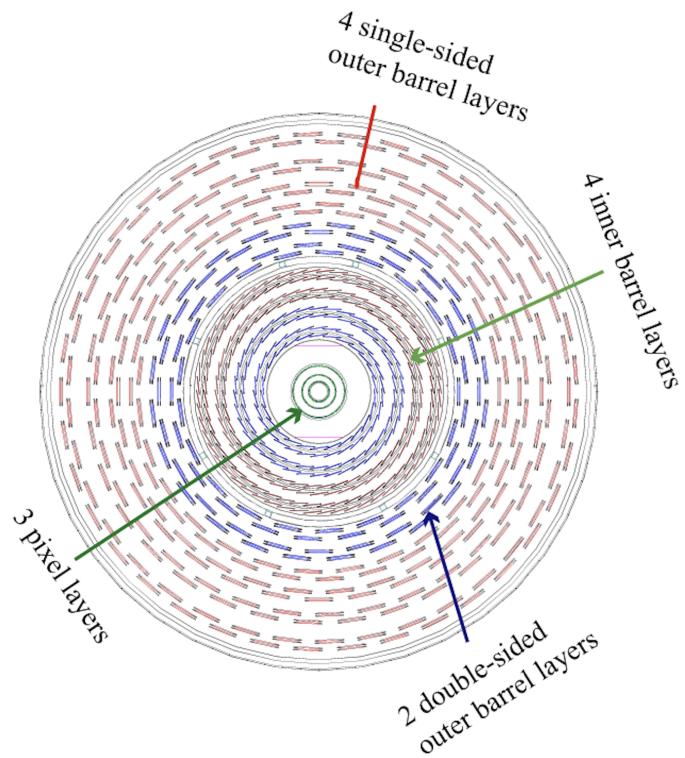


Figure 3-6. A transverse view of the silicon pixel and strip detectors, explicitly labelling the different layers involved.

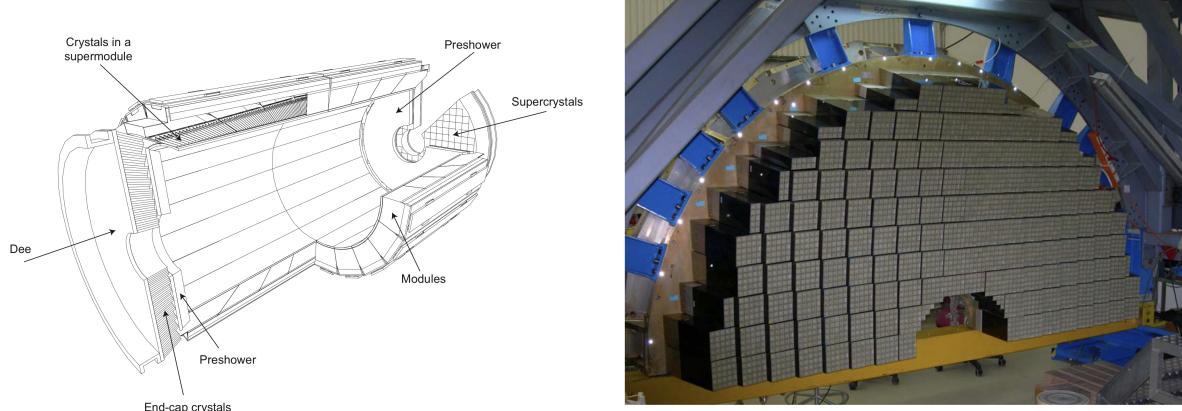


Figure 3-7. (Left) Cross-sectional view of the electromagnetic calorimeter of CMS. (Right) One of the Dees which comprise the EE. Each square of  $5 \times 5$  crystals constitutes a “super-crystal”. Figure taken from Ref. [10].



Figure 3-8. (Left) ECAL crystals made from PbWO<sub>4</sub> are grown in a lab. (Right) Although made mostly of metal, ECAL crystals are transparent and have a photomultiplier detector attached at the end.

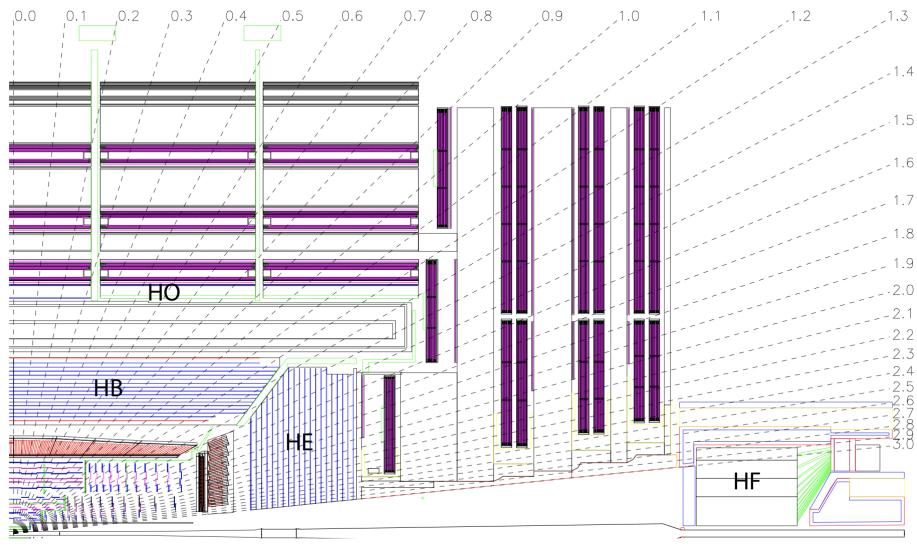


Figure 3-9. A cross-sectional quadrant view of CMS showing the locations of the HCAL components: the barrel (HB), outer (HO), endcap (HE), and forward (HF) detectors.

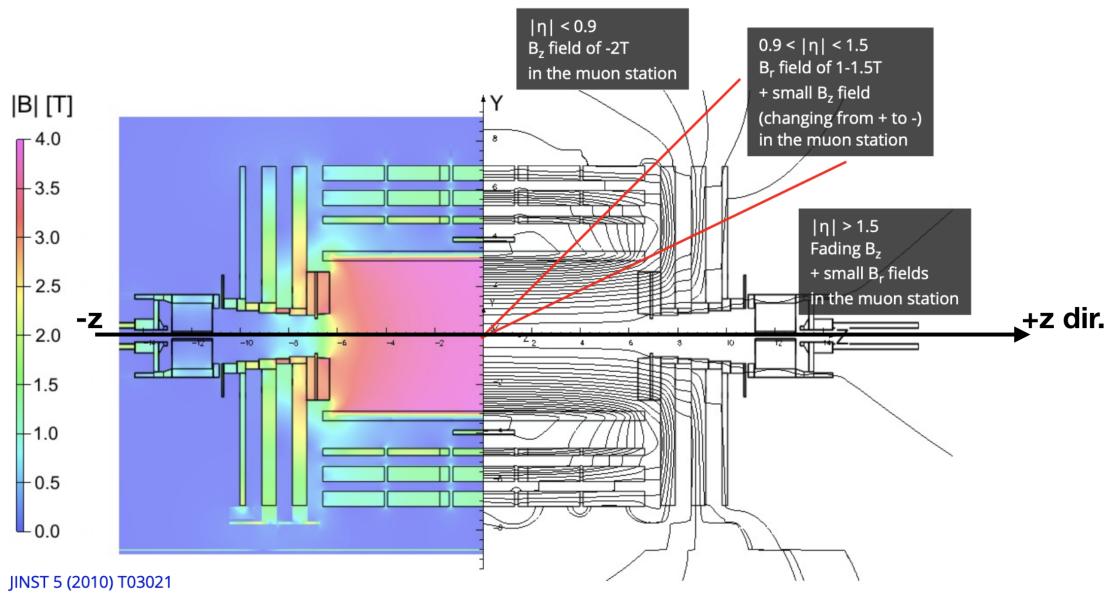


Figure 3-10. A longitudinal cross section of CMS showing the values of the magnetic field over the volume of CMS and various field lines. The magnetic field reaches its maximum of 3.8 T in the center of the detector.

# CHAPTER 4

## HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^* \rightarrow 4\ell$ CHANNEL

### 4.1 Motivation

The Higgs boson was discovered in 2012 by the CMS and ATLAS collaborations. This was a momentous achievement in particle physics because the existence of the Higgs boson was required to complete the SM. In fact, it is sometimes referred to as the “missing puzzle piece” of the SM. The Higgs boson is one of a kind: it is the only fundamental scalar particle ever discovered so far. The unique boson could be a portal to new physics (*beyond Standard Model physics*, BSM), e.g., by decaying into BSM low-mass dilepton mass resonances. In order to be certain that the recently discovered Higgs boson is truly the same as the one predicted by the SM, it is necessary to compare its measured properties to the predicted ones.

This chapter describes the Higgs boson mass measurement using full Run 2 data from the LHC. First, a general overview of the analysis workflow is given in Sec. 4.2. Afterwards, an analysis of the background estimation is given in Sec. 4.3.

### 4.2 Analysis Overview

The first step to performing a precision measurement of the Higgs boson mass ( $m_H$ ) is to “observe” many Higgs bosons. However, production of a Higgs boson is essentially nonexistent in everyday conditions and is still extremely rare even in the high-energy pp collisions of the LHC (Chapter 2). At a center-of-mass energy of 13 TeV, the total inclusive inelastic cross section of two protons colliding is 70 mb. Comparing this to the production cross section of a Higgs boson ( $\sigma_{pp \rightarrow H} = 59 \text{ pb}$ ) shows that a Higgs boson is produced in approximately one out of every *billion* pp collisions—a rare event indeed.

To complicate matters further, the Higgs boson has a *very* short mean lifetime of only  $1.6 \times 10^{-22} \text{ s}$  [9]. Thus, the Higgs boson is not directly detected by CMS (Chapter 3) but is instead *inferred* from its stable decay products that enter the various subdetectors. Among all the fundamental particles so far discovered, the Higgs boson bears the second heaviest mass (approximately 125 GeV), the first belonging to the top quark. This gives the scalar boson sufficient energy to decay into at least 9 different final states, where each decay occurs with a different probability—the *branching fraction* or *branching ratio* ( $\mathcal{B}$ )—whose value depends on

$m_H$ , as shown in Fig. 4-1. The question then becomes, “Which Higgs boson decay mode is most useful for the measurement of  $m_H$ ? ”.

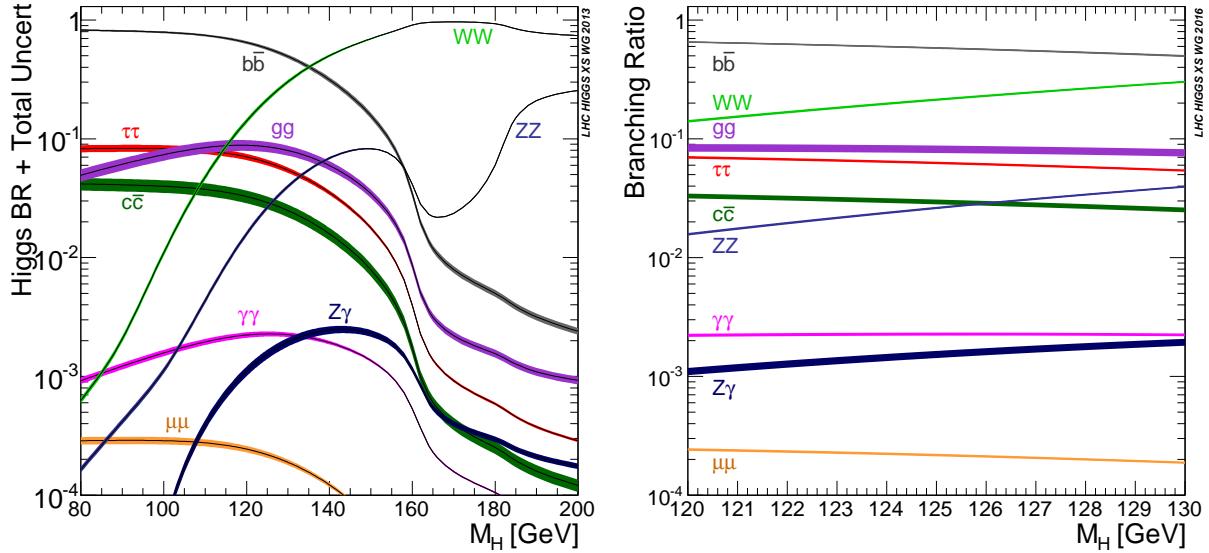


Figure 4-1. The branching ratios of various Higgs boson decays as a function of the Higgs boson mass over a wide range (Left) and a narrow range (Right) of values.

Owing to its large signal-to-background ratio of approximately 2 and its relatively rare four-lepton final state, the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay channel is chosen and is called the *signal process*. On average, a Higgs boson will decay into two Z bosons (one on-shell and one off-shell) only 2.6% of the time. In turn, each Z boson *may* decay into two opposite-sign, same flavor (OSSF) leptons ( $Z \rightarrow \ell^+ \ell^-$ , where  $\ell = e, \mu$ ) on average approximately 6.7% of the time. This signal process then gives rise to four distinct final states:  $4e$ ,  $4\mu$ ,  $2e2\mu$ ,  $2\mu2e$ . The branching ratio for the overall signal process is then calculated as:

$$\mathcal{B}(H \rightarrow ZZ^* \rightarrow 4\ell) = \mathcal{B}(H \rightarrow ZZ^*) [\mathcal{B}(Z \rightarrow \ell^+ \ell^-)]^2 = 1.8 \times 10^{-3}.$$

Thus, a signal event is expected to be produced only once in about every *trillion* pp collisions.

The strategy is then to search the pp collision data collected and analyzed by the CMS Experiment for all the detected  $H \rightarrow ZZ^* \rightarrow 4\ell$  events. The task is not so straightforward; events in the data are categorized—not by the entire decay process—but by their final state, based on

which triggers fired to collect which events. For each chosen event, the subdetectors of CMS (Secs. 3.1–3.3) provide a plethora of track and energy-detection information to reconstruct *objects*—representations of the underlying particles within the event. The reconstructed objects are then assembled in a fashion that checks if the logic coincides with the signal process of interest:  $H \rightarrow ZZ^* \rightarrow 4\ell$ . For example, a pair of OSSF lepton-like objects should appear to come from a Z-like object—i.e., having a nominal mass of approximately 91 GeV and zero net electric charge—instead of, say, appearing to come from a H-like object. Two such Z-like objects must be formed and should appear to come from a H-like object. All throughout, the reconstructed event must obey physics conservation laws (energy, momentum, charge, etc.) and the associated objects may even be required to pass certain detector selection criteria (e.g.,  $p_T^\mu > 5$  GeV). These criteria are analysis-specific and are collectively called the *event selection* of the analysis. The event selection for this analysis is described in Sec. ??.

Although the event selection is constructed to select only signal events, it is not guaranteed; there are certain physics process that have exactly the same initial and final states as the signal process. Such processes “contaminate” the collected signal events and are called *background processes*. Concretely, Fig. 4-2 shows how identical initial state gluons can react to produce exactly the same final state particles, while producing different intermediate particles: the signal process (Left), initiated by gluon-gluon fusion vs. a background process (Right) which skips the intermediate Higgs boson. It is imperative for all physics analyses to maximize the number of collected signal events while minimizing the number of collected background events. Section 4.3 discusses the associated background processes and how to estimate the number of events these contribute to the signal region.

Before drawing conclusions from the data themselves, it is necessary for particle physicists to make predictions about their analysis using simulated events or *simulation*. These events simulate a specific process (e.g.,  $pp \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$ ), governed by some theoretical framework that is programmed mathematically into a software package. Programs like `MADGRAPH5_amc@NLO` and `POWHEG` can simulate millions of rare (or even *fictitious*) events in

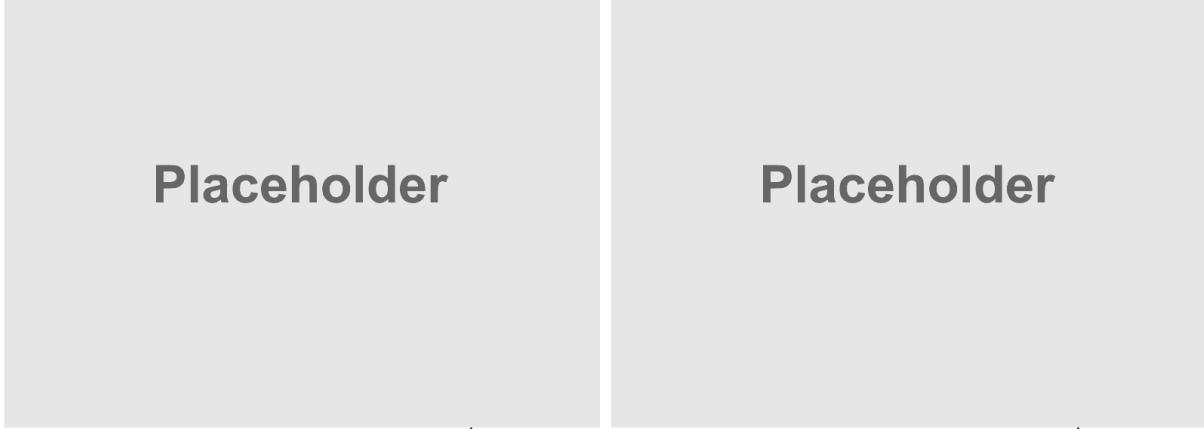


Figure 4-2. Feynman diagrams showing how the initial and final states are the same for the signal process ( $gg \rightarrow H \rightarrow ZZ^* \rightarrow 4\ell$ , Left) and one possible background process ( $gg \rightarrow ZZ^* \rightarrow 4\ell$ , Right).

just a single day, whereas the same event in actual data might otherwise take decades—or not at all! Furthermore, software can even simulate the particles as they travel through the simulated detectors. Programs like GEANT4 show analysts what to expect as the particles interact with a virtual version of the CMS detector. Predictions from simulation can then be compared to the truth—the data—as a way to check the accuracy of the analysis. For example, a surplus of events in data where none was expected may lead to the discovery of new particles, as was the case for the discovery of the Higgs boson.

So how is the measurement of  $m_H$  obtained? Since the signal process is  $H \rightarrow ZZ^* \rightarrow 4\ell$ , conservation of energy leads one to expect that  $m_{4\ell} \approx m_H$ . Although this is not how the final measurement is obtained, it is a logical starting point. The distribution of  $m_{4\ell}$  values reveals that the Higgs boson mass resonance stands well above the distribution of expected background events (Fig. 4-3). Simulated signal events are then used to predict the *line shape* of this signal peak. This signal modeling is performed using a double-sided Crystal Ball function to fit the line shape, for various mass points of  $m_H$ , in each of the four final states.

The precision of the measurement of  $m_H$  is improved by implementing several techniques: calculating a per-event matrix element kinematic discriminant, deriving correction factors for  $m_{4\ell}$  uncertainty for various regions of phase space, reevaluating the lepton  $p_T$  values using a  $Z_1$  mass

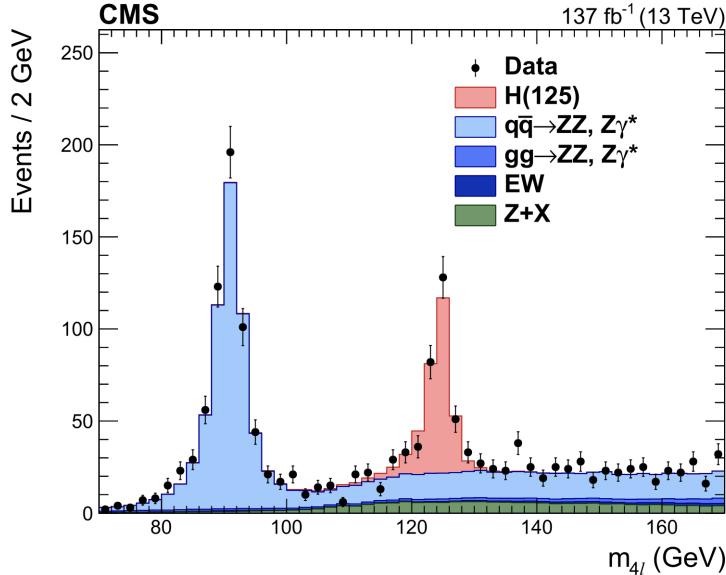


Figure 4-3. Distribution of  $m_{4\ell}$  from  $H \rightarrow ZZ^* \rightarrow 4\ell$  events using Full Run 2 data.

constraint, and constraining the four muon tracks to the selected vertex (*vertex constraint*).

Important in any scientific analysis is the careful study of all the associated uncertainties that inherently come with making *any* measurement. The two main kinds of uncertainties are either *statistical*, which depend on the number of data points used to make the measurement, or *systematic*, which depend on the instrumentation precision used to make the measurement. The systematic uncertainties for this analysis include:

- $\mathcal{L}_{\text{int}}$  (2.5%)
- Lepton identification and reconstruction efficiencies ( 2.5–9%)
- Lepton energy scale
- Estimation of the reducible background (40%).

The systematic theoretical uncertainties include:

- Renormalization and factorization scale uncertainties.
- Choice of the set of parton distribution functions.
- Branching fraction uncertainties for signal and background processes.

Finally, a likelihood fit is performed on the  $m_H$  spectrum to extract the most likely value of  $m_H$ .

### 4.3 Background Estimation

Measurement of the Higgs boson mass requires the accurate modeling of the total event yield in the signal region (SR), into which events can be categorized as either *signal* or *background*. These background events pass the signal event selection and thus spoil the purity of the signal events. This introduces further uncertainty into the final Higgs boson mass measurement. Therefore, it is a priority to reduce and to predict the expected number of background events, which can be split into two types: *irreducible background* (IB) and *reducible background* (RB) processes.

#### 4.3.1 Irreducible Background

IB processes produce two Z bosons and each Z subsequently decays into two prompt leptons (leptons that emerge directly from the primary vertex). This reliably produces a  $4\ell$  final state, whose prompt leptons are typically reconstructed as four leptons that pass tight selection (*PTS leptons*), as defined in Sections ?? and ?? . The IB event then looks indistinguishable from the 4 PTS lepton of the *signal* process and cannot be reduced; Thus, throwing away IB process could mean throwing away a signal event. Since IB events cannot be completely eliminated—or *reduced*—from the SR, they are given the name *irreducible* backgrounds. The two IBs for this analysis are:

- $gg \rightarrow ZZ \rightarrow 4\ell$  (gluon-gluon fusion),
- $q\bar{q} \rightarrow ZZ \rightarrow 4\ell$  (quark-antiquark annihilation).

#### 4.3.2 Reducible Background

While IB processes produce 4 prompt leptons, RB processes produce varying numbers of prompt and nonprompt leptons. Nonprompt leptons emerge from 3 main sources:

- misidentifying light-flavor hadrons (e.g.,  $\pi^\pm$ ) as leptons,
- heavy-flavor hadrons that decay mid-flight into leptons,
- and the asymmetric conversion of photons into electrons.

This analysis is tailored to efficiently reconstruct prompt(nonprompt) leptons as PTS(FTS) leptons. Ideally, the SR would contain only the events with 4 prompt leptons, which would always

be tagged as 4 PTS leptons. However, sometimes a truly nonprompt lepton is misidentified as a PTS lepton (sometimes called a *fake lepton*), depending on the kinematic properties of the lepton. This misidentification rate ( $f$ , sometimes called the *fake rate*) is due to imperfect detector performance, inefficiencies in reconstruction, and the specific choice of lepton selections used in the analysis.

If an event produces prompt and nonprompt leptons but is reconstructed as 4 PTS leptons (a 4P event), then it contaminates the SR. These processes constitute the RB processes (sometimes called “Z + X”). Examples of RB processes for this analysis include:

- Z + jets (yields 2 prompt leptons)
- $t\bar{t}$  + jets (yields 2 prompt leptons)
- WZ + jets (yields 3 prompt leptons)
- ZZ/ $Z\gamma^*$  + jets (yields 4 prompt leptons).

The careful estimation of these RB contributions to the 4P region is necessary for the precise measurement of the Higgs boson mass.

#### 4.3.2.1 OS Method

The goal of the OS Method is to estimate the number of opposite-sign same-flavor (OSSF)  $4\ell$  events produced by RB process that “contaminate” the 4P region ( $N_{4P}^{RB}$ ), given by Eq. 4-13. However, the typical approach of using simulated samples does not model RB well, since RB processes (e.g., Z + jets) rely on higher-order effects like jet modelling which are not yet accurately simulated. Instead, a data-driven approach is used.

The logic of the OS Method is to study events in data that are similar to, but not exactly the same as, those found in the 4P region. Thus, events in data are sorted into 2 control regions (CRs), both of which are orthogonal to the 4P region and to each other:

- the 3P1F CR (built from 3 PTS and 1 FTS leptons)
- the 2P2F CR (built from 2 PTS and 2 FTS leptons).

The event selection for the 2P2F and 3P1F CRs is almost identical to that of the SR, except that the FTS lepton(s) are required to build the  $Z_2$ . The events that contribute mostly to the

2P2F(3P1F) CR are those that produce 2(3) prompt and 2(1) nonprompt leptons and are called *2pr*(*3pr*) events. Similarly, the events that contribute mostly to the 4P SR are those that produce 4 prompt leptons and are called *4pr* events.

The final formula for  $N_{4P}^{\text{RB}}$  is obtained by first supposing that event  $k$  is a 2pr event and contributes to the 2P2F CR an event weight of  $w_{2\text{pr}\rightarrow 2\text{P2F}}^k$ . This weight is built from  $\hat{w}^k$ , which is the product of analysis weights (pileup, L1 pre-firing, etc.), and the reconstruction efficiencies ( $\epsilon$ ) of each lepton ( $\ell_n$ ):

$$w_{2\text{pr}\rightarrow 2\text{P2F}}^k = \hat{w}^k \cdot \epsilon_{\text{P}}^{\text{pr}}(\ell_1^k) \cdot \epsilon_{\text{P}}^{\text{pr}}(\ell_2^k) \cdot \epsilon_{\text{F}}^{\text{np}}(\ell_3^k) \cdot \epsilon_{\text{F}}^{\text{np}}(\ell_4^k), \quad (4-1)$$

where the superscript of  $\epsilon$  indicates the lepton promptness (pr = prompt, np = nonprompt), the subscript indicates the lepton tightness status (P = PTS, F = FTS). To simplify the equations that follow,  $\hat{w}^k$  is set to unity. If the reconstruction efficiencies of a particular category are the same for all  $j$  leptons across all  $k$  events (e.g.,  $\epsilon_{\text{P}}^{\text{pr}}(\ell_j^k) \equiv \epsilon_{\text{P}}^{\text{pr}}$ ), then Eq. 4-1 reduces to:

$$w_{2\text{pr}\rightarrow 2\text{P2F}}^k = \left( \epsilon_{\text{P}}^{\text{pr}} \right)^2 \left( \epsilon_{\text{F}}^{\text{np}} \right)^2. \quad (4-2)$$

Although a 2pr event mostly contributes to the 2P2F CR, a nonprompt lepton may be misidentified as a PTS lepton, depending on  $\epsilon_{\text{P}}^{\text{np}}$ . Such an event would then fall into the 3P1F CR and, allowing for only one nonprompt PTS lepton at a time, contributes an effective weight of

$$\begin{aligned} w_{2\text{pr}\rightarrow 3\text{P1F}}^k &= \left( \epsilon_{\text{P}}^{\text{pr}} \right)^2 \left[ \epsilon_{\text{P}}^{\text{np}}(\ell_1^k) \cdot \epsilon_{\text{F}}^{\text{np}}(\ell_2^k) + \epsilon_{\text{F}}^{\text{np}}(\ell_1^k) \cdot \epsilon_{\text{P}}^{\text{np}}(\ell_2^k) \right] \\ &= \left( \epsilon_{\text{P}}^{\text{pr}} \right)^2 [2\epsilon_{\text{P}}^{\text{np}}\epsilon_{\text{F}}^{\text{np}}]. \end{aligned} \quad (4-3)$$

Using the fact that a (non)prompt lepton is exclusively either PTS or FTS ( $\epsilon_{\text{P}}^{\text{pr(np)}} + \epsilon_{\text{F}}^{\text{pr(np)}} = 1$ ), while recognizing that  $\epsilon_{\text{P}}^{\text{np}} \equiv f$  (Sec. 4.3.2) and defining  $\epsilon_{\text{P}}^{\text{pr}} \equiv \epsilon$ , allows Eq. 4-3 to be written as

$$w_{2\text{pr}\rightarrow 3\text{P1F}}^k = 2\epsilon^2 f(1-f). \quad (4-4)$$

Even more rarely, both nonprompt leptons from a 2pr event may be misidentified as PTS

leptons. In this case, event  $k$  contributes to the 4P region an effective weight of

$$w_{2\text{pr} \rightarrow 4\text{P}}^k = \epsilon^2 f^2, \quad (4-5)$$

where it is assumed that both leptons have the same misidentification rate. Similar equations can be derived for the contributions of a 3pr event to the 3P1F CR and to the 4P region:

$$w_{3\text{pr} \rightarrow 3\text{P1F}}^k = \epsilon^3 (1 - f) \quad (4-6)$$

$$w_{3\text{pr} \rightarrow 4\text{P}}^k = \epsilon^3 f. \quad (4-7)$$

Since a 3pr event needs only 1 nonprompt PTS lepton to be included in the 4P region (therefore, carrying only 1 factor of  $f$ ), a 3pr event tends to contribute more weight to 4P than does a 2pr event (which carries  $f^2$ ).

If the total number of 2pr, 3pr, and 4pr events is  $X_{2\text{pr}}$ ,  $X_{3\text{pr}}$ , and  $X_{4\text{pr}}$ , respectively, then Fig. 4-4 shows how the weight of a single event, derived in the previous equations (and others forthcoming), from each category contributes to the final yield of each CR ( $N_{2\text{P2F}}$ ,  $N_{3\text{P1F}}$ ) and of the SR ( $N_{4\text{P}}$ ). It is then straightforward to evaluate the expected number of RB 4P events:

$$\begin{aligned} N_{4\text{P}}^{\text{RB}} &= \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 4\text{P}}^k + \sum_{m=1}^{X_{3\text{pr}}} w_{3\text{pr} \rightarrow 4\text{P}}^m \\ &= \sum_{k=1}^{X_{2\text{pr}}} \epsilon^2 f^2 + \sum_{m=1}^{X_{3\text{pr}}} \epsilon^3 f \\ &= f^2 \epsilon^2 X_{2\text{pr}} + f \epsilon^3 X_{3\text{pr}}, \end{aligned} \quad (4-8)$$

where only the quantities  $f$  and  $\epsilon$  are known, so  $X_{2\text{pr}}$  and  $X_{3\text{pr}}$  must be estimated. This is achieved

by relating  $X_{2\text{pr}}$  to  $N_{2\text{P}2\text{F}}$  using Eq. 4-2 across all 2pr events:

$$\begin{aligned}
 N_{2\text{P}2\text{F}} &= \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 2\text{P}2\text{F}}^k \\
 &= \sum_{k=1}^{X_{2\text{pr}}} \epsilon^2 (1-f)^2 \\
 &= (1-f)^2 \epsilon^2 X_{2\text{pr}}
 \end{aligned} \tag{4-9}$$

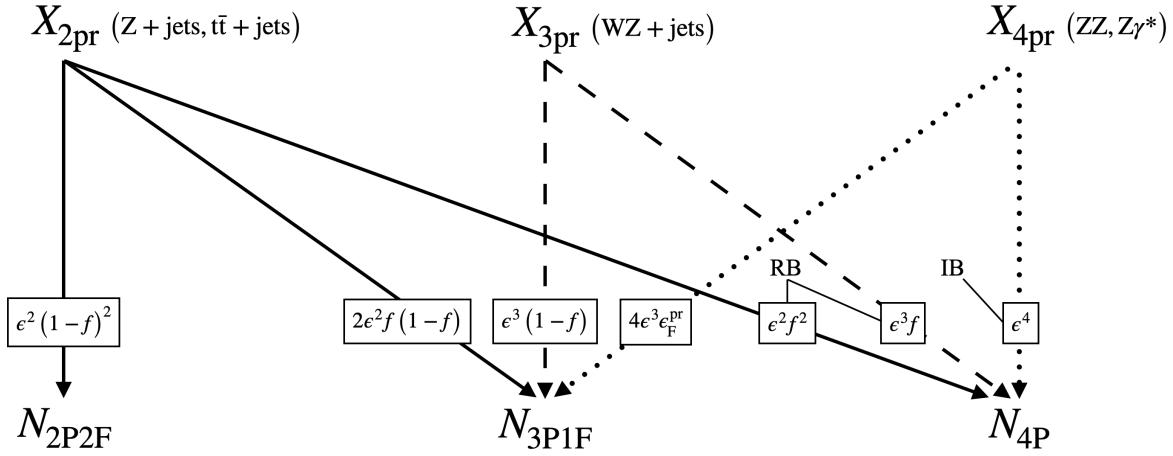


Figure 4-4. Contributions of per-event weights (boxed values) of various  $n$ -prompt-lepton processes ( $X_{n\text{pr}}$ , in parentheses) to the total numbers of events in the observed control regions ( $N_{2\text{P}2\text{F}}, N_{3\text{P}1\text{F}}$ ) and signal region ( $N_{4\text{P}}$ ). The labels “RB” and “IB” indicate those contributions which comprise the reducible and irreducible backgrounds, respectively.

The strategy to relate  $X_{3\text{pr}}$  to  $N_{3\text{P}1\text{F}}$  is not as straightforward as relating  $X_{2\text{pr}}$  to  $N_{2\text{P}2\text{F}}$ , since two other sources also contribute to the 3P1F CR (as shown in Fig. 4-4):

- A 2pr RB process can yield one nonprompt PTS lepton (via Eq. 4-4).
- A 4-prompt IB process ( $q\bar{q}/gg \rightarrow ZZ \rightarrow 4\ell$ , “ZZ”) can yield one FTS lepton.

The second of these is well estimated from simulation, since ZZ produces 4 prompt leptons. If the total number of simulated ZZ events is  $X_{4\text{pr}}^{\text{ZZ}}$ , then event  $k$  belonging to these events contributes to the 3P1F CR an effective weight of

$$w_{4\text{pr}, \text{ZZ} \rightarrow 3\text{P}1\text{F}}^k = 4\epsilon^3 \epsilon_F^{\text{pr}}, \tag{4-10}$$

which accounts for any of the 4 prompt leptons to be reconstructed as a FTS lepton while the others are PTS leptons. Incorporating Eqs. 4-4, 4-6, and 4-10 for all events relates  $X_{3\text{pr}}$  to  $N_{3\text{P1F}}$ :

$$\begin{aligned}
N_{3\text{P1F}} &= \sum_{j=1}^{X_{3\text{pr}}} w_{3\text{pr} \rightarrow 3\text{P1F}}^j + \sum_{k=1}^{X_{2\text{pr}}} w_{2\text{pr} \rightarrow 3\text{P1F}}^k + \sum_{m=1}^{X_{4\text{pr}}^{\text{ZZ}}} w_{4\text{pr}, \text{ZZ} \rightarrow 3\text{P1F}}^m \\
&= \sum_{j=1}^{X_{3\text{pr}}} \epsilon^3 (1-f) + \sum_{k=1}^{X_{2\text{pr}}} 2\epsilon^2 f (1-f) + \sum_{m=1}^{X_{4\text{pr}}^{\text{ZZ}}} 4\epsilon^3 \epsilon_{\text{F}}^{\text{pr}} \\
&= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + 4\epsilon^3 \epsilon_{\text{F}}^{\text{pr}} X_{4\text{pr}}^{\text{ZZ}} \\
&= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + N_{3\text{P1F}}^{\text{ZZ}}, \tag{4-11}
\end{aligned}$$

where  $N_{3\text{P1F}}^{\text{ZZ}}$  is simply the raw (integer) number of ZZ events that pass 3P1F selections, obtained directly from simulation.

At this point,  $N_{4\text{P}}^{\text{RB}}$  can be isolated by combining Eqs. 4-8, 4-9, and 4-11:

$$N_{4\text{P}}^{\text{RB}} = \left( \frac{f}{1-f} \right) N_{3\text{P1F}}^{\text{Data}} - \left( \frac{f}{1-f} \right)^2 N_{2\text{P2F}}^{\text{Data}} - \left( \frac{f}{1-f} \right) N_{3\text{P1F}}^{\text{ZZ}}. \tag{4-12}$$

Thus, Eq. 4-12 estimates the RB contribution to the 4P region by using a single lepton misidentification rate to reweight the raw yields of events found in the 3P1F and 2P2F CRs of data, and also the 3P1F CR of ZZ.

It should be mentioned that the above formula was derived assuming that the per event analysis weights were set to unity ( $\hat{w}^k = 1$ ) before being scaled by the misidentification rates. It was also assumed that  $f$  is a constant, for all nonprompt leptons misidentified as PTS leptons across all events, which is not the case as is shown in Figures 4-13 and 4-14. Thus, an extension of Eq. 4-12 can be formed by assigning misidentification rates that depend on the kinematical variables per lepton, by restoring the analysis weights per event, and by summing over the total number of raw yields per CR:

$$N_{4\text{P}}^{\text{RB}} = \sum_{i=1}^{N_{3\text{P1F}}^{\text{Data}}} \hat{w}_i^i \frac{f_i}{1-f_i} - \sum_{j=1}^{N_{2\text{P2F}}^{\text{Data}}} \hat{w}_j^j \frac{f_{1,j}}{1-f_{1,j}} \frac{f_{2,j}}{1-f_{2,j}} - \sum_{k=1}^{N_{3\text{P1F}}^{\text{ZZ}}} \hat{w}_{\text{ZZ}}^k \frac{f_k}{1-f_k}, \tag{4-13}$$

where  $f_{1,j}$  and  $f_{2,j}$  are the misidentification rates of the first and second FTS leptons, respectively,

found in the  $j^{\text{th}}$  2P2F event and  $\hat{w}_{ZZ}^k$  accounts for the differential QCD and electroweak  $K$  factors.

The prediction of Eq. 4-4 can be seen in Figs. 4-9–4-12 (red line) for 2016–2018 UL data.

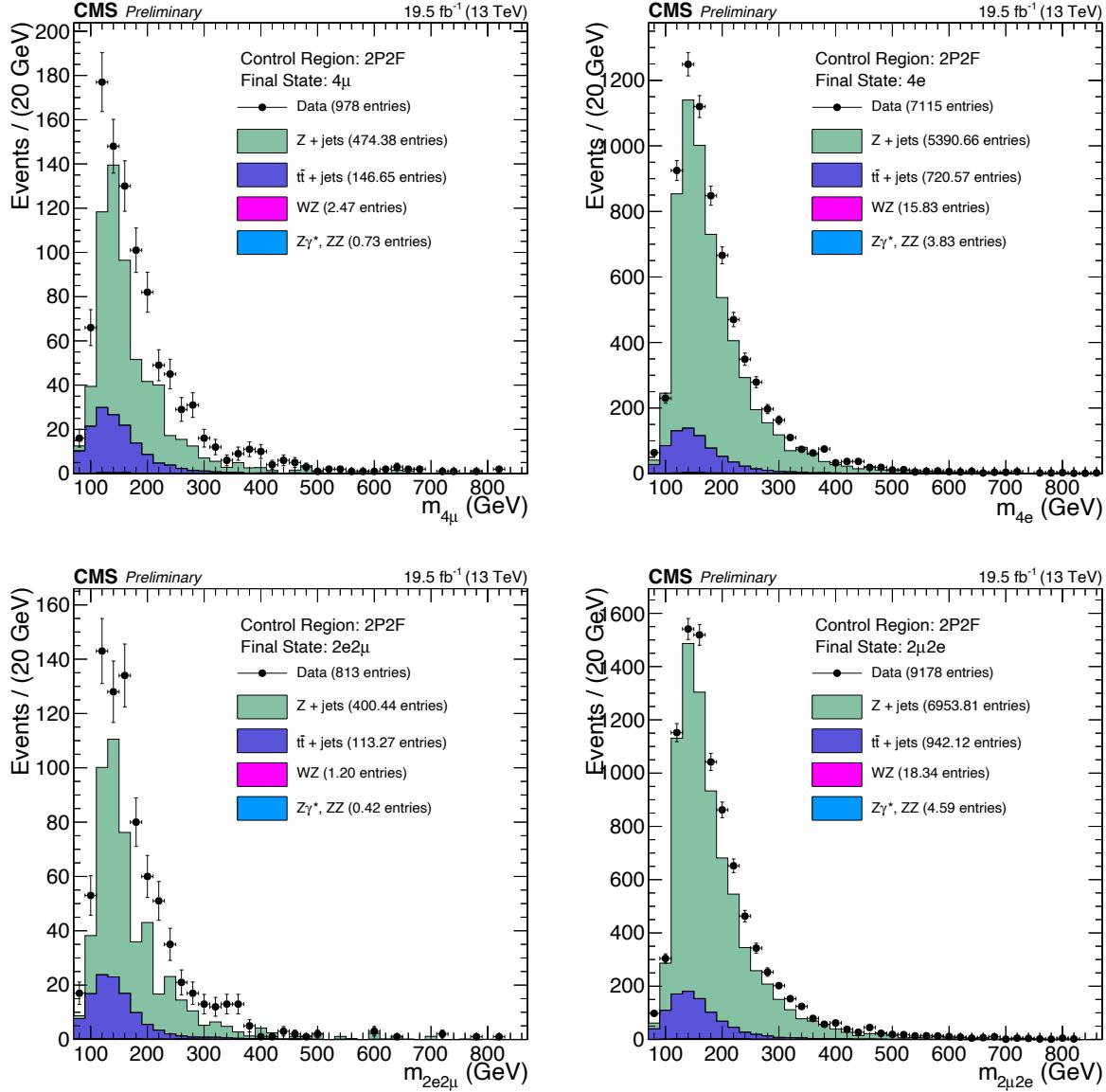


Figure 4-5. Events from 2016 pre-VFP UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ( $Z + \text{jets}$ ,  $t\bar{t} + \text{jets}$ ,  $WZ$ ,  $ZZ$ ,  $Z\gamma^*$ ). The results are split into the  $4\ell$  final states:  $4\mu$  (top left),  $4e$  (top right),  $2e2\mu$  (bottom left),  $2\mu2e$  (bottom right).

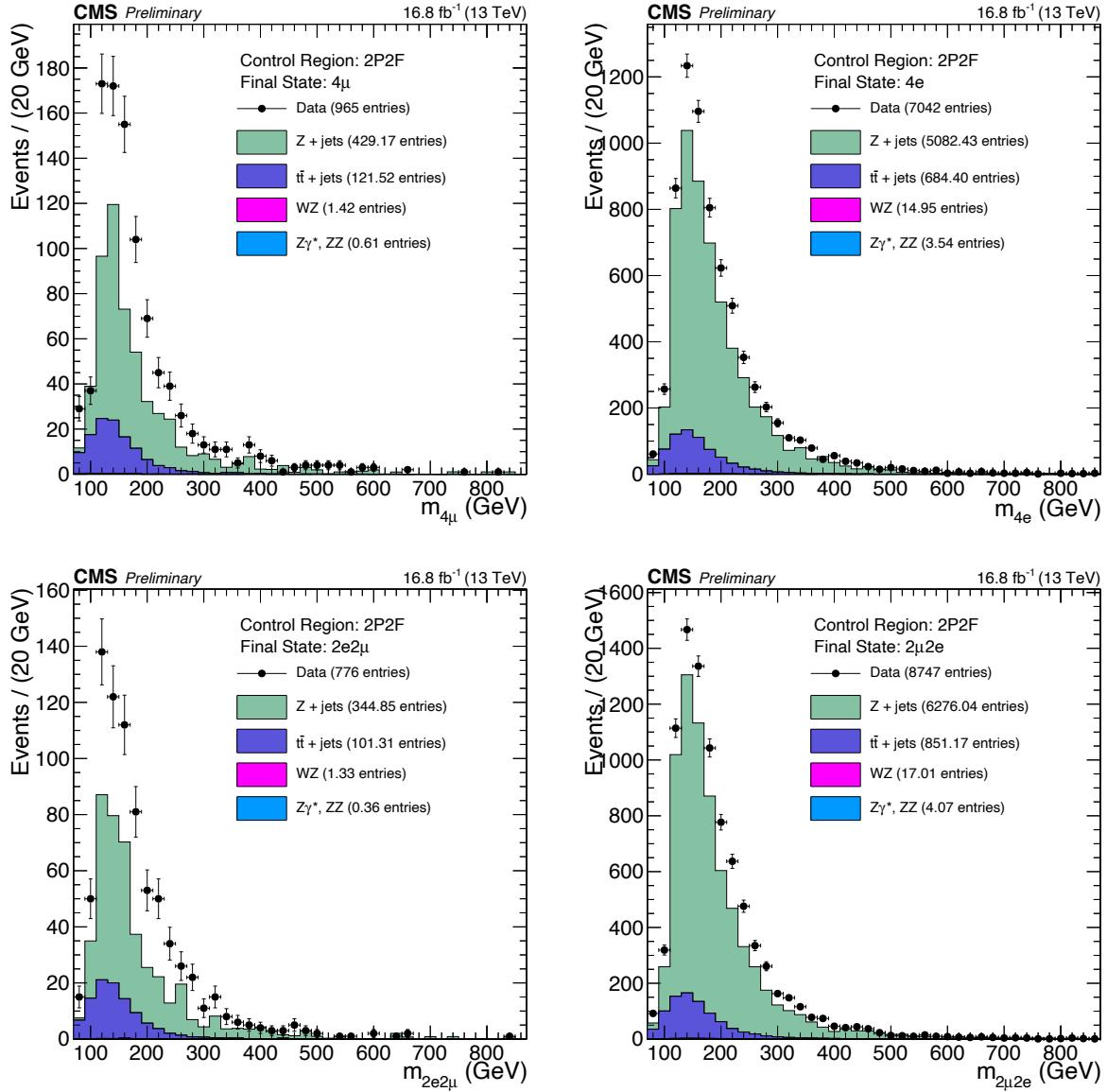


Figure 4-6. Events from 2016 post-VFP UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ( $Z + \text{jets}$ ,  $t\bar{t} + \text{jets}$ , WZ, ZZ,  $Z\gamma^*$ ). The results are split into the  $4\ell$  final states:  $4\mu$  (top left),  $4e$  (top right),  $2e2\mu$  (bottom left),  $2\mu2e$  (bottom right).

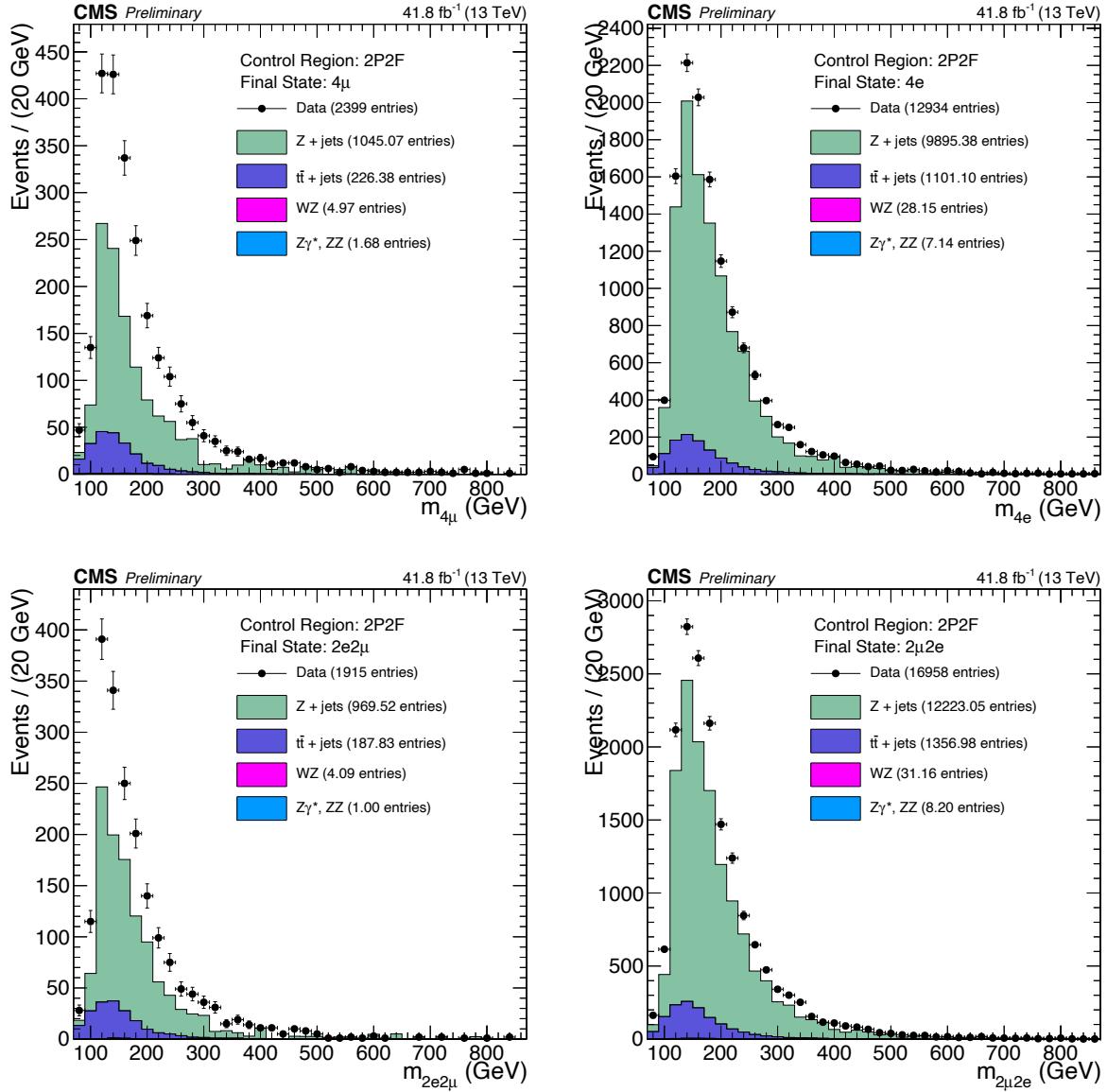


Figure 4-7. Events from 2017 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples (Z + jets, t̄t + jets, WZ, ZZ, Z $\gamma^*$ ). The results are split into the 4 $\ell$  final states: 4 $\mu$  (top left), 4e (top right), 2e2 $\mu$  (bottom left), 2 $\mu$ 2e (bottom right).

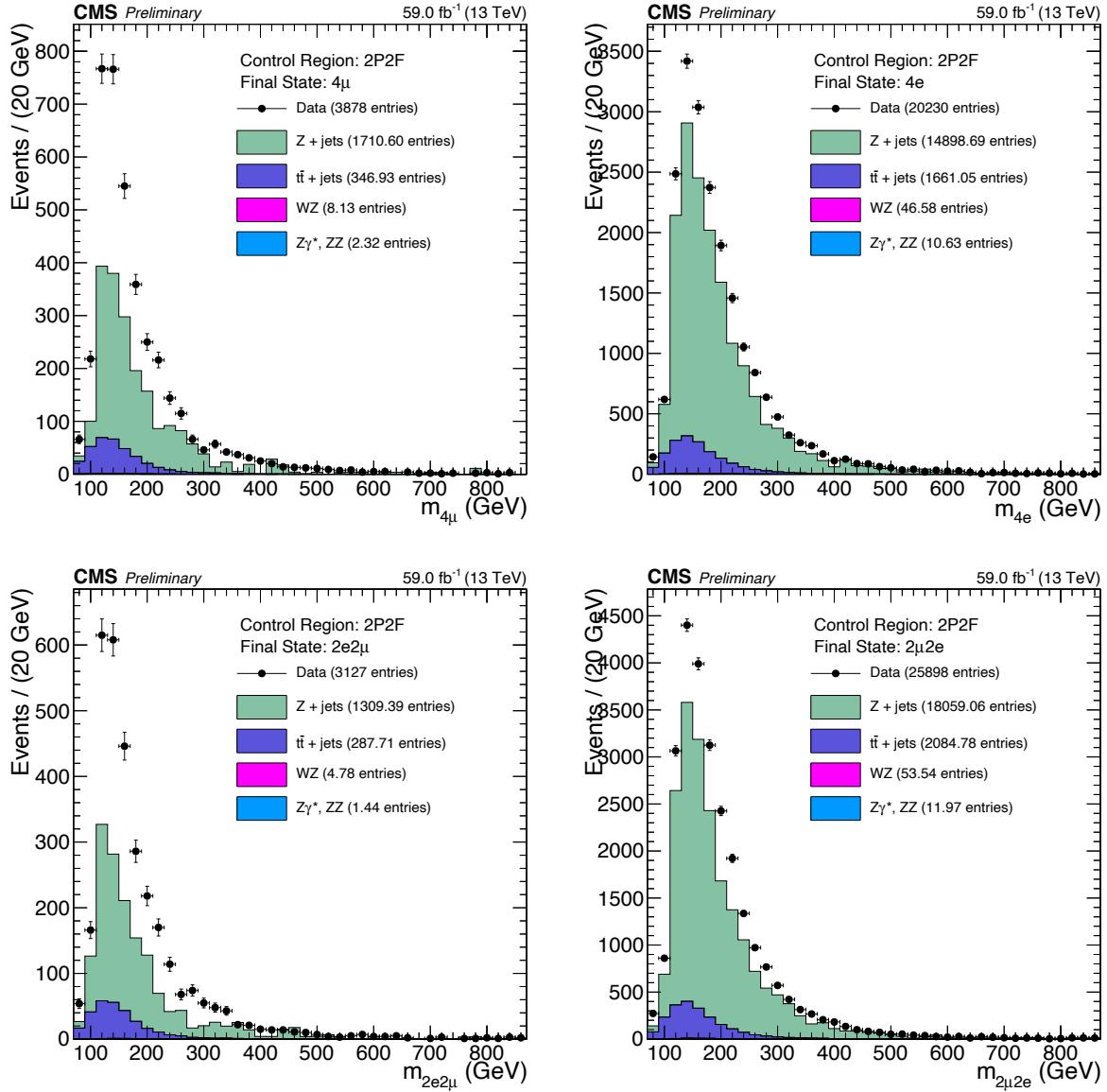


Figure 4-8. Events from 2018 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ( $Z + \text{jets}$ ,  $t\bar{t} + \text{jets}$ , WZ, ZZ,  $Z\gamma^*$ ). The results are split into the  $4\ell$  final states:  $4\mu$  (top left),  $4e$  (top right),  $2e2\mu$  (bottom left),  $2\mu2e$  (bottom right).

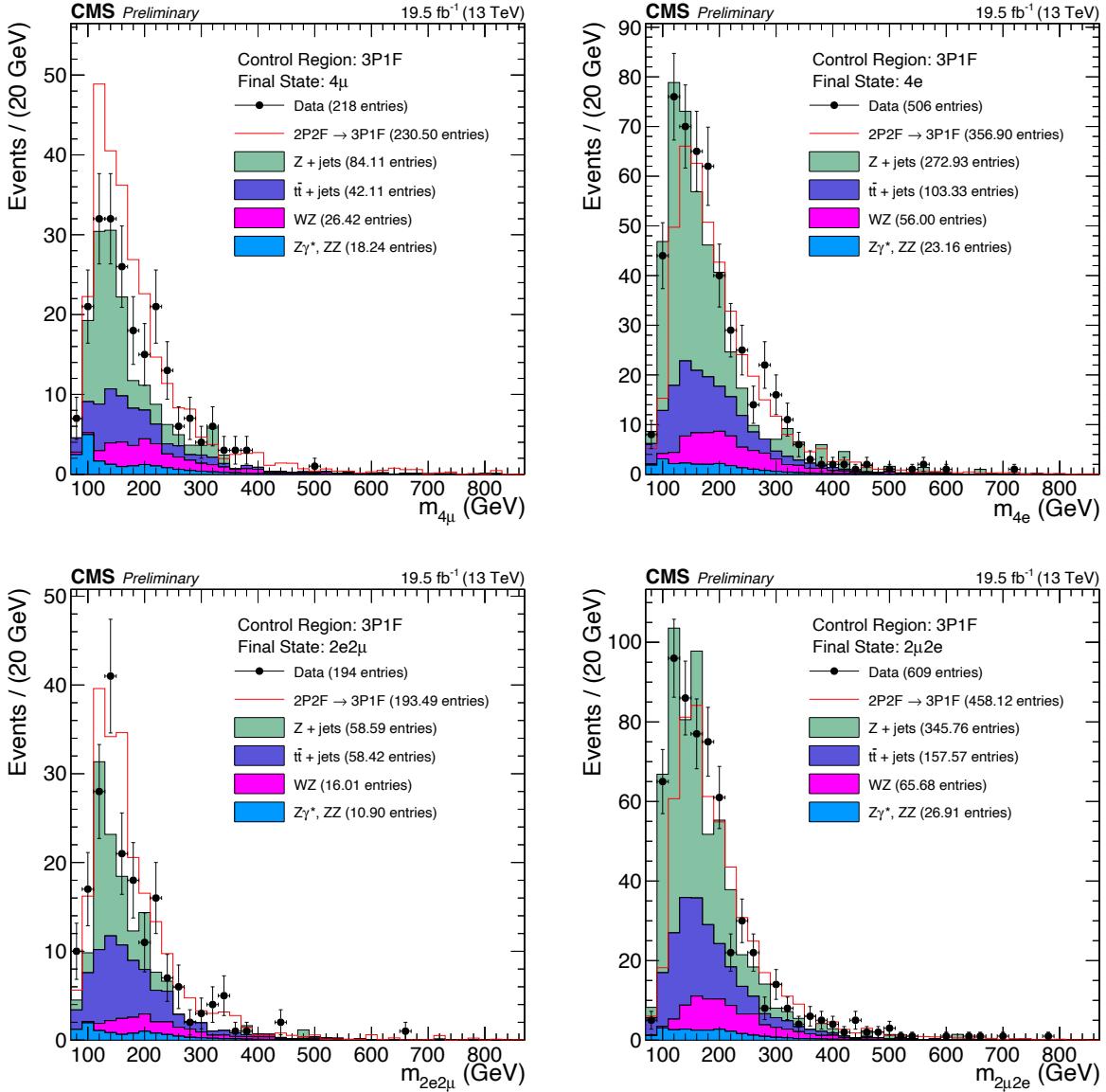


Figure 4-9. Events from 2016 pre-VFP UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples (Z + jets, t̄t + jets, WZ, ZZ, Z $\gamma^*$ ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 4-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4 $\ell$  final states: 4 $\mu$  (top left), 4e (top right), 2e2 $\mu$  (bottom left), 2 $\mu$ 2e (bottom right).

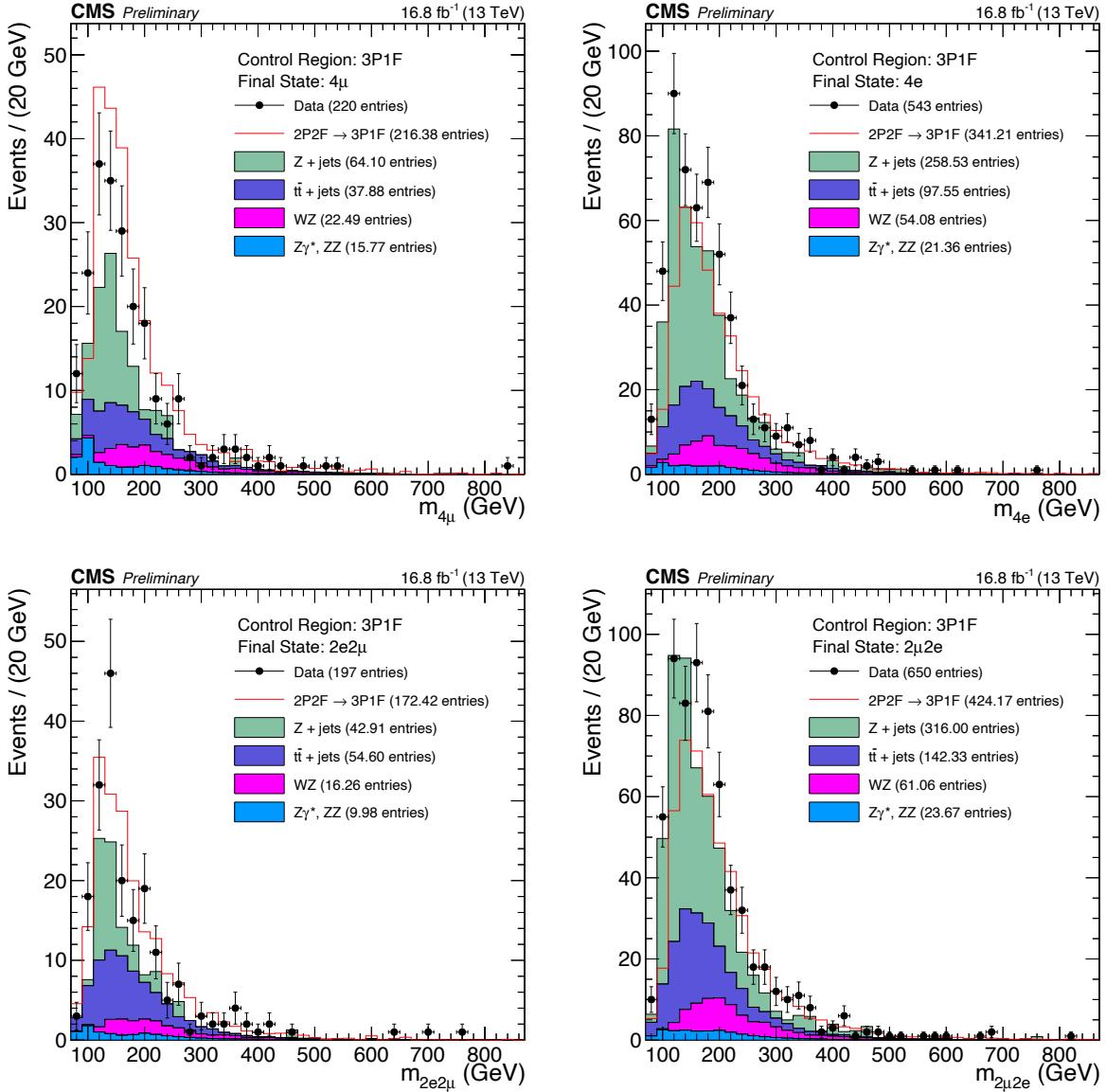


Figure 4-10. Events from 2016 post-VFP UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ( $Z + \text{jets}$ ,  $t\bar{t} + \text{jets}$ ,  $WZ$ ,  $ZZ$ ,  $Z\gamma^*$ ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 4-4 and stacking that on top of the  $WZ$  and  $ZZ$  distributions. The results are split into the  $4\ell$  final states:  $4\mu$  (top left),  $4e$  (top right),  $2e2\mu$  (bottom left),  $2\mu2e$  (bottom right).

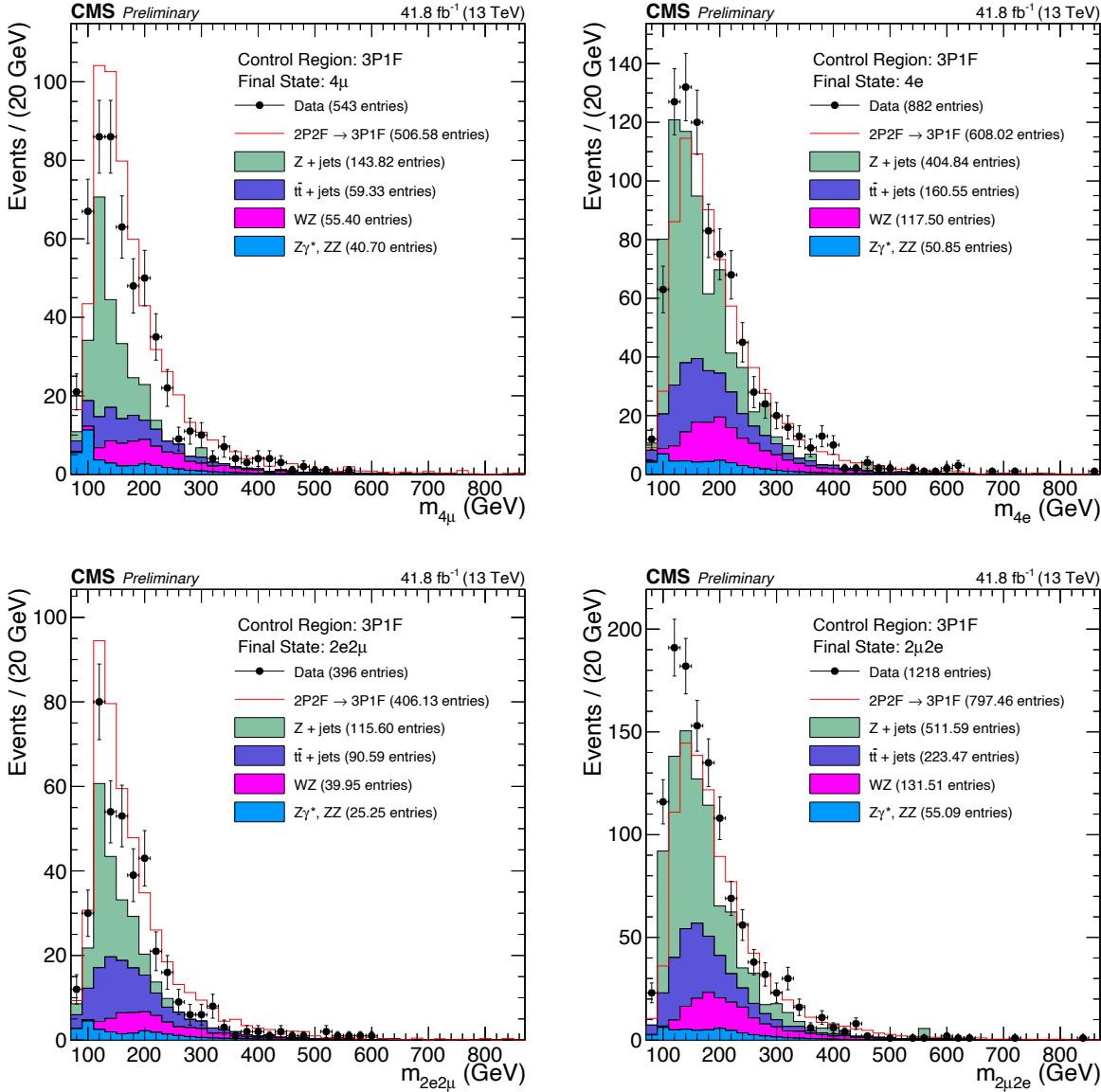


Figure 4-11. Events from 2017 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ( $Z + \text{jets}$ ,  $t\bar{t} + \text{jets}$ ,  $WZ$ ,  $ZZ$ ,  $Z\gamma^*$ ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 4-4 and stacking that on top of the  $WZ$  and  $ZZ$  distributions. The results are split into the  $4\ell$  final states:  $4\mu$  (top left),  $4e$  (top right),  $2e2\mu$  (bottom left),  $2\mu2e$  (bottom right).

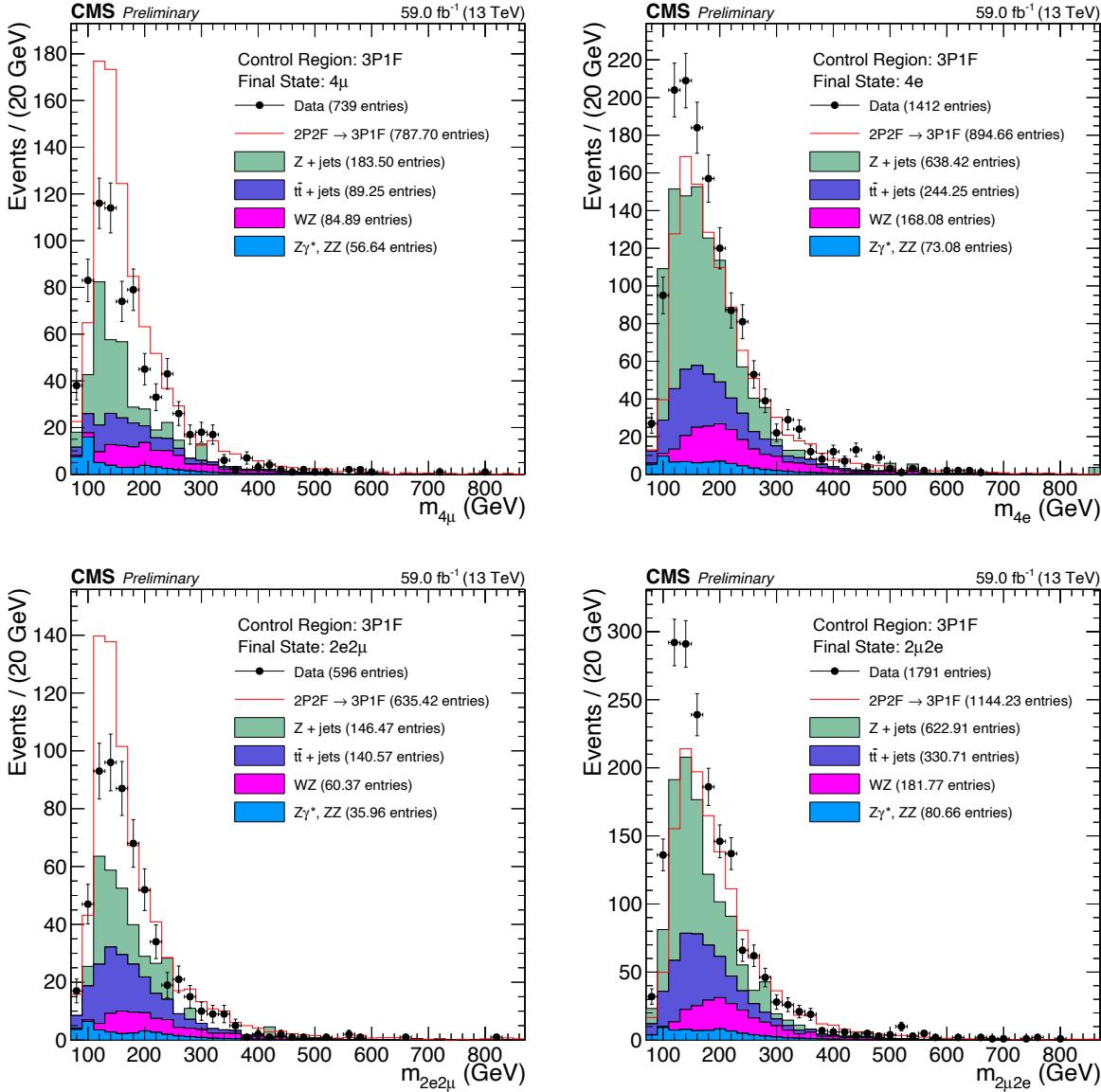


Figure 4-12. Events from 2018 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples ( $Z + \text{jets}$ ,  $t\bar{t} + \text{jets}$ , WZ, ZZ,  $Z\gamma^*$ ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 4-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the  $4\ell$  final states:  $4\mu$  (top left),  $4e$  (top right),  $2e2\mu$  (bottom left),  $2\mu2e$  (bottom right).

### 4.3.2.2 Lepton misidentification rate measurement

As mentioned in Sec. 4.3.2, the lepton misidentification rate ( $f$ ) is the probability that a nonprompt lepton will pass tight selections (PTS). The value  $f$  is a function of the flavor ( $\ell = e, \mu$ ),  $p_T$ , and  $\eta$  of a lepton. The misidentification rate is calculated by simply counting the number of nonprompt PTS leptons ( $N_P^{\text{np}}$ ) that enter a particular  $\ell, p_T, |\eta|$  bin compared to the total number of loose probe leptons ( $N_L^{\text{np}}$ ) in the same bin:

$$f(\ell, |\eta|, p_T) = \frac{N_P^{\text{np}}}{N_L^{\text{np}}}. \quad (4-14)$$

The  $p_T^e$  bin edges are [5–10–20–30–40–50–80] GeV and the  $p_T^\mu$  bin edges are [5–7–10–20–30–40–50–80] GeV. The nonprompt leptons used to measure  $f$  are taken from events in data with a signature like that of  $Z + \ell_L$ , where  $Z$  is a  $Z$  boson and  $\ell_L$  is a loose lepton ( $e, \mu$ ). By construction, this region of events is orthogonal to the 2P2F, 3P1F, and 4P regions, and provides a clean source of  $\ell_L$ . The loose lepton, whose selection is defined in Sections ?? and ??, is also called the *probe* lepton. The probe lepton is either a PTS or FTS lepton and is counted towards both the numerator and denominator of Eq. 4-14.

Events are selected that satisfy the following criteria:

- The event has exactly 3 leptons.
- The event contains  $E_T^{\text{miss}} < 25$  GeV.
- Two of the leptons form a  $Z$  candidate. A  $Z$  candidate is formed when:
  - The lepton pair is OSSF.
  - Both leptons PTS.
  - The leading lepton has  $p_T > 20$  GeV.
  - The subleading lepton has  $p_T > 10$  GeV.
  - The lepton pair satisfies  $|m_{\ell\ell} - m_{Z_{\text{PDG}}}| < 7$  GeV.
- The third and final lepton is loose (i.e., either a PTS or FTS lepton).
- Suppress QCD processes: probe lepton and OS lepton from  $Z$  have  $m_{\ell\ell} > 4$  GeV.

The calculation of  $f$  requires that  $\ell_L$  is a nonprompt lepton but since data events were used, this may not be the case. For example, the decay of  $WZ$  produces 3 prompt leptons and so this

contribution must be subtracted. Thus, the number of expected prompt probe leptons from WZ events is subtracted from both the numerator and denominator in Eq. 4-14 for each  $\ell, p_T, |\eta|$  bin. The final misidentification rates used in the OS Method for electrons and muons are shown in Figures 4-13 and 4-14 using 2016–2018 UL data.

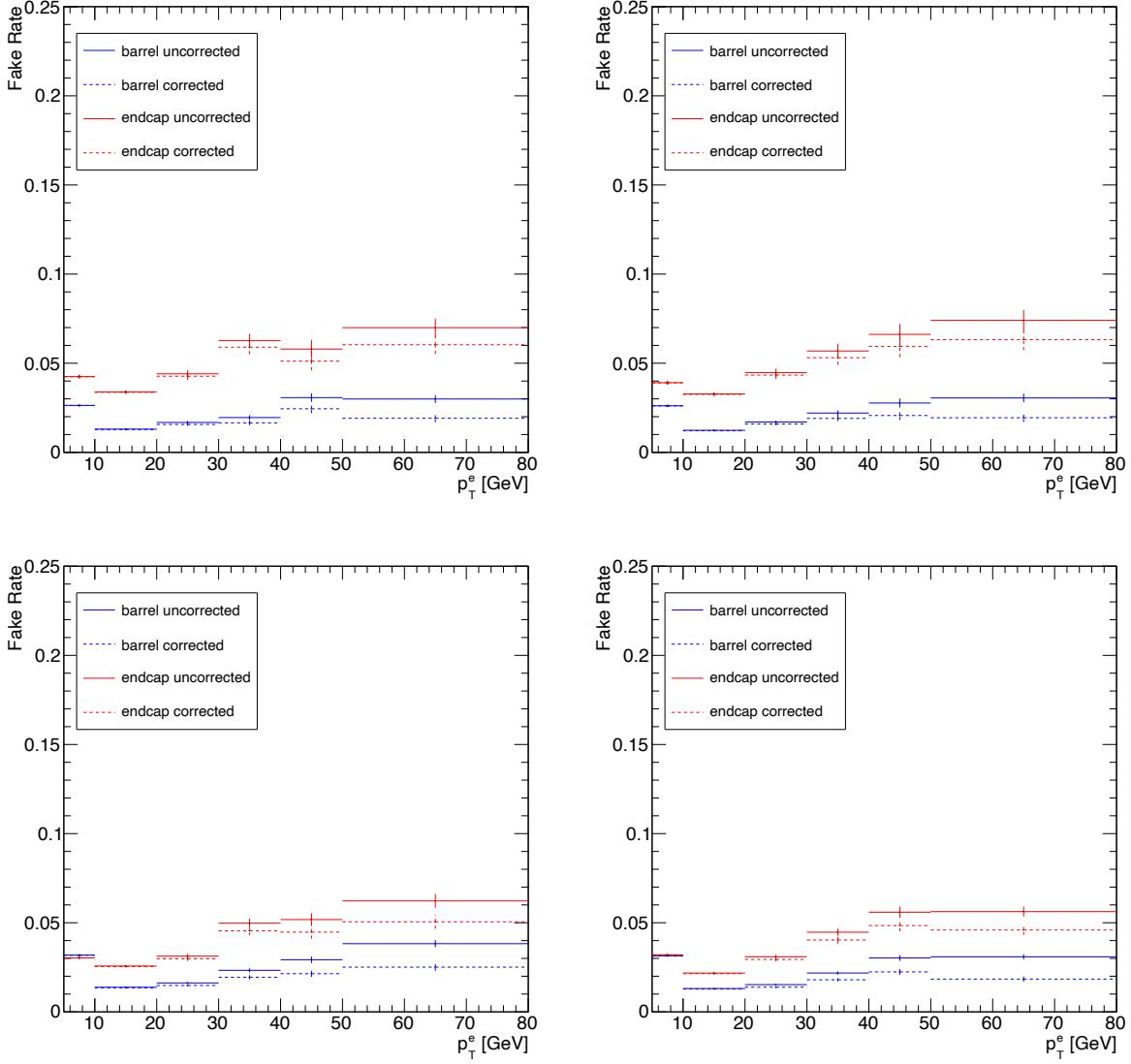


Figure 4-13. Electron misidentification rates vs. the  $p_T$  of the probe electron evaluated for the OS Method selecting  $Z + e$  events in UL data for each year in Run 2: 2016 pre-VFP (top left), 2016 post-VFP (top right), 2017 (bottom left), and 2018 (bottom right). Rates are evaluated separately for the barrel (blue lines) and endcap (red lines) regions, partitioned at  $|\eta^e| = 1.497$ . The final rates (dashed lines) are calculated by subtracting out the expected  $WZ$  contribution—estimated from simulation—from the numerator and denominator in the rate calculation.

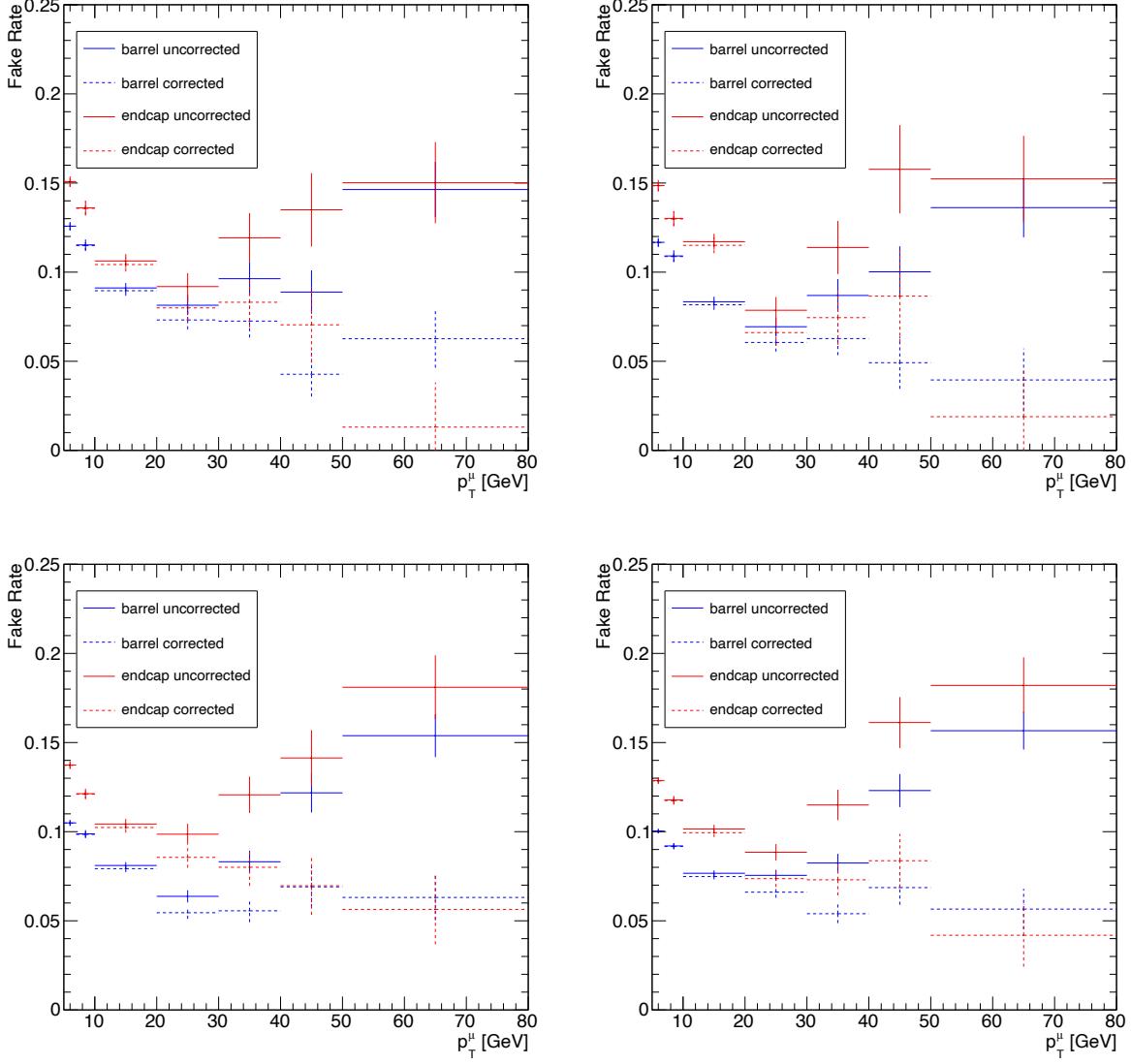


Figure 4-14. Muon misidentification rates vs. the  $p_T$  of the probe muon evaluated for the OS Method selecting  $Z + \mu$  events in UL data for each year in Run 2: 2016 pre-VFP (top left), 2016 post-VFP (top right), 2017 (bottom left), and 2018 (bottom right). Rates are evaluated separately for the barrel (blue lines) and endcap (red lines) regions, partitioned at  $|\eta^\mu| = 1.2$ . The final rates (dashed lines) are calculated by subtracting out the expected  $WZ$  contribution—estimated from simulation—from the numerator and denominator in the rate calculation.

## CHAPTER 5 CONCLUSION

Using data collected by the CMS detector, the mass of the Higgs boson was measured to be  $m_H = 125.38 \pm 0.11 \text{ GeV}$ ; the world's most precise measurement to date. This measurement studied the  $H \rightarrow ZZ^* \rightarrow 4\ell$  ( $\ell = e, \mu$ ) decay channel and used LHC Run 2 data from pp collisions, corresponding to an integrated luminosity of  $137.1 \text{ fb}^{-1}$ . The analysis improved upon previous measurements of  $m_H$  by constraining muon tracks to the primary vertex (*vertex constraint*).

## REFERENCES

- [1] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Physics Letters B* **716** (2012) 30.
- [2] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer et al., *Observation of a new boson with mass near 125 GeV in pp collisions at  $\sqrt{s} = 7$  and 8 TeV*, *Journal of High Energy Physics* **2013** (2013) 81.
- [3] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, A.A. Abdelalim et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012) 1.
- [4] ATLAS COLLABORATION AND CMS collaboration, *Combined measurement of the Higgs boson mass in pp collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS experiments*, *Phys. Rev. Lett.* **114** (2015) 191803.
- [5] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, T. Bergauer, M. Dragicevic et al., *A measurement of the Higgs boson mass in the diphoton decay channel*, *Physics Letters B* **805** (2020) 135425.
- [6] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. Abidi et al., *Measurement of the Higgs boson mass in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  channels with  $\sqrt{s} = 13$  TeV pp collisions using the ATLAS detector*, *Physics Letters B* **784** (2018) 345.
- [7] A. Andreassen, W. Frost and M.D. Schwartz, *Scale-invariant instantons and the complete lifetime of the standard model*, *Phys. Rev. D* **97** (2018) 056006.
- [8] CDF collaboration, *High-precision measurement of the W boson mass with the CDF II detector*, *Science* **376** (2022) 170.
- [9] Particle Data Group, P.A. Zyla, R.M. Barnett, J. Beringer, O. Dahl, D.A. Dwyer et al., *Review of Particle Physics*, *Progress of Theoretical and Experimental Physics* **2020** (2020) 083C01.
- [10] CMS collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08004.

## BIOGRAPHICAL SKETCH

Jake Rosenzweig had the best childhood anyone could ask for, growing up in Jacksonville, FL: enjoying video games with excellent friends, playing football on the beach, and having plenty of opportunity to make mistakes. He graduated from the University of Florida in 2011 with a B.S. in chemistry, while maintaining his sanity by getting minors in education and Latin. He enjoys building things from scrap, weightlifting, hiking in the Coloradoan mountains, gardening, silence, and—most of all—receiving the beleaguered stare from his wife after telling her a *particularly* bad dad joke.